Journal of
Clinical Psychology

APRIL 1951
Volume VII Number 2

Contents

A Method for Objective Quantification of Certain Countertransference Attitudes. Fred E. Fiedler, University of Chicago 101
An Evaluation of the Use of Difference Scores in Prediction. J. R. Wittenborn, Yale University 108
Catinua, Catemensions, Catescales. Frank M. Du Mas, University of Texas 112
A Scale of Neuroticism: An Adaptation of the Minnesota Multiphasic Personality Inventory. John F. Winne, Veterans Administration Hospital, Northampton, Mass. 117
Brief Tests of Intelligence in the Psychiatric Clinic. John R. Knott, Richard G. Cannicott, John P. Umberger and Ina McD. Bilodeau, University of Iowa 123
The Immediate Test. Raymond Corsini, Wisconsin Dept. of Public Welfare 127
The Hillside Short Form of the Wechsler-Bellevue. Milton S. Gurvitz, Hillside Hospital 131
Wechsler-Bellevue Subtest Scatter in the Affective Disorders. Samuel Waldfogel and William Guy, McLean Hospital 135
A Test-Retest Evaluation of the Wechsler Forms I and II with Mental Defectives. William Hays and Bernard Schneider, Lincoln State School and Colony 140
A Study of the Differential Responses on the Vocabulary Sub-Test of the Wechsler-Bellevue Intelligence Scale. Chalmers L. Stacey and Bernard Portnoy, Syracuse University 144
The Effect of Laterality Localization of Focal Brain Lesions on the Wechsler-Bellevue Subtests. A. Lloyd Andersen, Nevada State Department of Health 149
The Stereognostic Capacity of Brain Injured As Compared With Familial Mental Defectives. William Sloan and Gerard J. Bensberg, Lincoln State School and Colony 154
A Comparison of Recall and Recognition Types of Measurement on Verbal Items, and Their Implications for Deterioration Testing. E. E. Glik, Fitzsimons General Hospital 157
A Comparison of the Stimulus Value of Szondi's Pictures With Those of Normal Americans. Wilson H. Guertin, Michigan State College. 163
The Letter Series as a Projective Technique. Helen M. Wiltshire, Kent State University 167
Clinical Psychology in Public Schools. Douglas Courtney, Ohio State University 171
Case Study of an Attitudinal Pathosis. Frederick C. Thorne, University of Vermont 176

(Continued on next page)

JOURNAL OF CLINICAL PSYCHOLOGY

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Medical College, University of Vermont

HOWARD F. HUNT
University of Chicago

JERRY W. CARTER, JR.
U. S. Public Health Service

WILLIAM A. HUNT
Northwestern University

ELAINE F. KINDER
Rockland (N. Y.) State Hospital

SAMUEL W. HAMILTON

ROBERT I. WATSON
Washington University

C. M. LOUTTIT
University of Illinois

ELEANOR C. THORNE, *Business Manager*

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science. Special sections in each issue will be devoted to news, comments, and letters to the editors on matters pertinent to the profession. Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

ORIGINAL ARTICLES are published only with the understanding that they are contributed exclusively to the *Journal of Clinical Psychology*. The publishers are not responsible for statements made or opinions expressed by contributors in articles printed in these columns. Copyrights cover publication in the *Journal of Clinical Psychology* and articles may not be reproduced without permission of the publishers.

MANUSCRIPTS must be typewritten, double-spaced with good margins, on one side of the paper only. The publishers will not be responsible for manuscripts lost in the mails or through unavoidable accident. Retain a carbon copy after sending the original to the editor. It is requested that all articles be as brief as consistent with clarity. All manuscripts are subject to editorial modification.

THE STYLE of all manuscripts including bibliographies must be in accordance with the instructions given in: Anderson, J. E., and Valentine, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.*, 1944, 41, 345-376. References should be arranged alphabetically according to the directions given above. Full address of the author should appear somewhere in the article, preferably at the end.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscribers' expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Burlington, Vermont, by Frederick C. Thorne, M.D. Subscription price $6.50; Canadian subscriptions $7.00; foreign subscriptions $7.50 including postage; single copies $2.00. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1951, Frederick C. Thorne, M. D.

Contents

(*Concluded from Front Cover*)

Parent-Child Similarities in Personality Disturbances. E. Lakin Phillips, Arlington (Va.) Guidance Center 188

The Constancy of the Altitude: A Note. Leonard Diller and Robert M. Beechley, Northern New Jersey Mental Hygiene Clinic 191

A Note on the Limitations of the Use of the Bender Gestalt Test as a Diagnostic Aid in Patients with a Functional Complaint. Leo J. Hanvik, Colorado A. & M. College 194

Writing Psychological Reports. Austin Foster, University of Texas ... 195

Editorial Opinion 196

Books Received 198

JOURNAL OF
CLINICAL PSYCHOLOGY

VOL. VII APRIL 1951 No. 2

A METHOD OF OBJECTIVE QUANTIFICATION OF CERTAIN COUNTERTRANSFERANCE ATTITUDES

FRED E. FIEDLER*

The University of Chicago

THE PROBLEM

The importance of the countertransference has been stressed by many theorists in the field of psychotherapy. (4, pp. 85-114; 6, pp. 53-54) Freud defines it as the feeling "which arises in the physician as a result of the patient's influence on his (the therapist's) unconscious feelings," (2, p. 289) which has here been interpreted to mean the feelings extraneous to the treatment proper which the therapist wittingly and unwittingly holds toward his patient. The present paper will describe one method of objective quantification of such attitudes.

THE RATIONALE OF THE METHOD

The therapist's feelings, perceptions, and expectations which are unrealistic, are countertransference phenomena. If the perception of the therapist is distorted, we must assume that the distortion is caused by unconscious feelings within the therapist and not by other sources in the treatment situation. If we, therefore, obtain a measure of this misperception we also will obtain an estimate of the countertransference feelings which are involved. Some of these misperceptions most likely arise when the therapist views the patient in relation to himself. Wittingly or unwittingly he probably asks himself at one time or another questions such as, "How different is this patient from me?" "Is this patient's problem somewhat like mine?" The therapist who unwarrantedly perceives his patient to be much more like himself has clearly a quite different feeling toward him than the therapist who sees his patient as unjustifiedly different from himself. The patient may thus be viewed as a screen upon which the therapist projects his countertransference feelings. By separating the reality of the situation from the therapist's perceptions we can then estimate the nature and the intensity of the countertransference.

AN EXPLORATORY STUDY

While we do not at present possess any method by which we can obtain an objective picture of how the patient "really "is, we can ask the patient to describe himself. This is not as satisfactory as a "true personality picture" would be, but can be utilized for the purposes of the method to be described below. This has been done in the context of an exploratory study which gives an example of the procedure as well as the results which may be expected. However, because of the smallness of the sample as well as a variety of uncontrolled factors in the design, these results must be interpreted with caution.

*The present paper grew out of research of the VA Research Project on the Selection of Clinical Psychologists. The writer acknowledges with pleasure the stimulation and assistance which he received from his co-workers, Dr. D. W. Fiske, Mr. K. S. Isaacs, and Miss Sarah Counts. Thanks are also due o Drs. J. G. Miller, Hedda Bolgar, W. D. Neff, and H. F. Hunt for their critical evaluations and uggestions which they contributed at various stages of this paper.

Twenty-two psychotherapists, mostly clinical psychologists working in a mental hygiene clinic with psychoneurotic veterans, were asked to cooperate in this study. The method involved the following steps:

1. The patient was asked to describe himself by means of an array of seventy-six statements. These statements were drawn from Murray's *Explorations in Personality* (3, pp. 142-242) and contain statements indicative of the needs and presses postulated by Murray's theory. The patient was instructed to sort the seventy-six statements into eight categories of 1, 5, 12, 20, 20, 12, 5, and 1 statement each, ranking them from the most descriptive statements at one extreme of the distribution to the statement least descriptive at the other extreme.

2. The therapist was requested to predict how the patient described himself, using the same statements and the same instructions, (i. e., "try to duplicate the patient's self-descriptions").

3. The therapist was also asked (a) to describe himself with these statements and (b) to describe how he would ideally like to be.

4. For each sorting, every statement is assigned a score from 1 to 8 corresponding to the category into which it was placed. Each sorting is thus represented by a distribution of scores, and the entire distribution can be correlated with any other distribution made up by the traits or statements which were used. In this case, the four sortings, the patient's self-description, the therapist's prediction, the therapist's self-description and the therapist's "ideals" were intercorrelated. This statistical method follows Stephenson's Q technique (5).

As the object of prediction we have chosen the self-description of the patient because it represents a complex and challenging predictive task for the clinician. Moreover, the patient's self-description, i.e. his public self-concept, is an important factor in his psychological make-up which the therapist may reasonably be expected to know. In intercorrelating the set of four descriptions, we obtain four correlations pertinent to this method:

1. The *real similarity* (*RS*) between the therapist's and the patient's self-sorts.

2. The *assumed similarity* (*AS*), that is, the similarity which the therapist assumes to exist between himself and his patient. This is obtained by correlating the therapist's self-sort with his prediction of the patient.

3. The *real similarity to the therapist's ideal* (*RSI*) which is obtained by correlating the therapist's ideal with the patient's self-sort.

4. The *assumed similarity to the therapist's ideal* (*ASI*) which is obtained by correlating the therapist's ideal with the prediction which he made of the patient.

We shall be concerned only with two scores derived from these correlations since the measures taken by themselves are partly dependent on factors not directly related to countertransference. (Because a therapist's and his patient's self-sorts may actually resemble one another very closely, the therapist would be justified in assuming a high degree of similarity between his own and the patient's sorts.) We are therefore interested in the amount to which he over- or underestimates the patient's similarity to himself or the patient's similarity to the therapist's ideal. Inasmuch as correlations are non-linear measures, they cannot be treated like ordinary test scores. Hence the correlations are squared. The squared "real similarity" is subtracted from the squared "assumed similarity" to obtain the amount of *unwarranted assumed similarity* (UAS = $r^2_{AS} - r^2_{RS}$). Similarly $r^2_{ASI} - r^2_{RSI}$ gives us the amount of *unwarranted assumed similarity to the ideal* (UASI).

It must be remembered that we are here concerned with obtaining an estimate of the clinician's over- or underestimation of the similarity in self-sorts. Squaring minus correlations would eliminate all minus signs, and thus distort the estimates. The signs must therefore be regarded as directional and as such should be retained. For example, a therapist's and his patient's self-sorts correlate −.20; the therapist's prediction of the patient's self-sort and the therapist's self-sort correlate +.10, $r_{AS}^2 = .0100$, and $r_{RS}^2 = .0400$. If we subtract r^2_{AS} from r^2_{RS} we obtain a difference of −.0300. This is clearly misleading since the therapist estimated the patient to be more like himself than was warranted by the real similarity[1](−.20). If we retain the signs, however, we find that the therapist overestimated his parient's similarity to himself by .0500 ($r^2_{AS} - R^2_{RS} = .100 - (-.0400) = .0500$.) UAS or UASI may then be positive or negative depending on whether the therapist assumes the similarity between himself (or his ideal) to be greater or less than the real similarities would warrant.

The interpretation of this over- or underestimation which seems most meaningful to this writer (though by no means the only possible one) is as follows: The clinician's overestimation of the patient's similarity to himself, that is, predicting the patient's sort as more similar to his own than it is in fact, might mean that the therapist wants the patient to be more like himself than the patient really is. Perceiving someone as more like oneself probably means accepting him as a person like oneself, one whose feelings one can understand: a generally benign or empathic attitude. Conversely, seeing the patient as less like oneself indicates a pulling away from the patient, a dissociating the patient from oneself: a generally negative attitude.

If the therapist assumes the patient to be more like his ideal than the patient's self-sort indicates, this might mean that he unconsciously perceives the patient as better adjusted than the patient is. This will lead him to expect more from the patient than the patient is capable of giving, and may be described as a demanding attitude on the part of the therapist. (Such unconscious demands may be that the patient live up to the therapist's standards, or that the patient be better adjusted and therefore give love, affection, and gratitude to the therapist.) On the other hand, seeing the patient as much less like his ideal (i.e. less well adjusted and self-sufficient) would suggest that the therapist sees the patient as a relatively more helpless individual, one to whom he must give affection and succorance since he, the therapist, is by contrast much more healthy and able. In the extreme it might mean seeing the patient as a completely worthless person. In general, then, UAS has been interpreted as the therapist's expression of positive or negative feelings toward the patient, the zero point indicating an attitude of objectivity. UASI is seen as an index of the therapist's demandingness or succorance toward the patient, the zero point being an undemanding attitude.

If we tentatively accept these interpretations of UAS as falling on the dimensions of empathic or positive-negative feelings, and UASI as a demanding-supporting dimension we shall expect on the basis of theoretical and empirical knowledge about therapy that the following patterns will be found:

a. Good therapists will not manifest negative feelings toward their patients: UAS will be positive or at least neutral.

b. Good therapists will like their patients more (that is, give more of themselves) than they will demand from their patients in return. In other words, provided UAS is better than neutral, UASI should be less than UAS.

Table 1 presents the twenty-five possible combinations of over- and underestimation of UAS and UASI with interpretations which are tentatively advanced by the writer as an educated guess about their meaning.

TABLE 1. TENTATIVE INTERPRETATIONS OF COUNTERTRANSFERENCE PATTERNS.

UAS	UASI: ++	+	O	−	−−
++	overwhelming patient with kindness, high empathy	genuine interest in patient's prompt recovery	very warm and undemanding attitude	succorant and helpful	mothering, infantilizing patient
+	therapist's need for narcissistic gains predominates	patient seen as instrument to therapist's satisfaction	warmth and acceptance	supportive, helpful	succorant
O	narcissistic, calculating, cold	somewhat demanding, aloof	objective, professional attitude	helpful in a professional manner	aloof, do-good spirit
−	Rejecting and impatient with patient's progress	rejecting and demanding	cold and aloof	only intellectually interested, cold curiosity	grudging helpfulness
−−	punitively demanding without willingness to help patient	very hostile and demanding	very hostile and cold	rejecting but overtly trying to help	very hostile and deprecating

UAS is interpreted here as falling on a continuum of benign, empathetic, positive attitudes to hostile, cold, rejecting attitudes.

UASI is interpreted as indicating the emotional demands which the therapist makes of his patient, versus the emotional support he is ready to give.

++ indicated over-estimation of .0900 or above
+ indicates over-estimation of .0200—.0899
O indicates over-estimation of −.0199 to +.0199
− indicates under-estimation of −.0200—.0899
−− indicates under-estimation of −.0900 or below

Results

Table 2 presents a rank-ordered list of therapists rated by a very responsible therapy supervisor who has had extensive contacts with 16 of the 22 therapists, and has discussed their treatment cases with them over the space of one to four years. Six of the therapists, who had not been supervised by the first supervisor were ranked by asking their own supervisors to compare them with therapists in the list. This is not too satisfactory a method, but serves for this particular purpose. As can readily be seen, the very competent group shows UAS and UASI patterns which are theoretically very desirable while the poorer group has generally less desirable patterns.

Table 2. Countertransference Patterns of 22 Ranked Therapists

Competence	*Therapist*	*Rank Order*	*UAS*	*UASI*	*UAS*[1]	*UASI*
Very competent	A	1	.2288[1]	.0952[1]	++	++
Very competent	B	2.5	.1105	−.0360	++	—
Very competent	C	2.5	.0936	−.0056	++	0
Very competent	D	4.5	.0297	.0096	+	0
Very competent	E	4.5	.0627	−.0112	+	0
Very competent	F	6	.0000	.0000	0	0
Very competent	G	7	.2537	−.0675	++	—
Very competent	H	8	.0205	−.0279	0	—
Very competent	I	9	−.0195	.0320	0	+
Very competent	J	10	.0000	−.0275	0	—
Very competent	K	11	.0056	−.0280	0	—
Competent	L	12	.0209	.0037	0	0
Competent	M	13	.0255	−.0125	+	0
Competent	N	14	.0775	.0493	+	+
Poor	O	15.5	−.0425	−.1472	—	— —
Poor	P	15.5	.0324	−.0315	+	—
Poor, but exceptionally successful case	Q	17	.0735	.0540	+	+
Poor	R	18	−.0425	−.4800	—	— —
Poor	S	19	.0040	.0345	0	+
Poor	T	20	−.0232	−.0040	—	0
Poor	U	21	.0147	−.1189	0	— —
Poor	V	22	.0255	.0130	—	0

rho
Therapeutic rating/UAS = 59 (1% level)
Therapeutic rating/UASI = 08
Therapeutic rating/UAS-UASI = 26

[1] 0 is assigned to r^2 of $\pm$ .0225 (significantly different from zero correlation at about 20% level)
+ or — is assigned to r^2 of $\pm$ .0226–.0899 (significantly different from zero correlation at about 1% level)
++ or — — is assigned to r^2 of $\pm$ .0900

The following additional checks are presented as evidence that the interpretations here advanced are meaningful and useful concepts, though they are by no means the only possible solutions. As the interpretations in Table 1 suggest, a high positive UAS indicates a benign and supportive therapist; conversely, a negative UAS, indicates a very hostile therapist. The rank-order correlation between ranked therapeutic competence and UAS yields a rho of .59, significant below the 1 per cent level.

It is, of course, very difficult to validate descriptions of unconscious attitudes. One further check involved asking one well trained psychologist to listen to recorded treatment interviews by six therapists for whom both recordings and countertransference patterns were available. (Both kinds of data, recordings and predictions, were available for only three of the six patients, however.) Since there are five poss-

ible steps for UAS and five for UASI, she was asked to identify which of these five steps were applicable to each of the six therapists. She was able to predict the patterns for two therapists, and she predicted UAS for two additional therapists. That is, she predicted six of the possible twelve steps correctly when the chance prediction for each was 1/5. This yields a P of .016, that is, significant below the 2 per cent level. Actually, this prediction does not tell the entire story since our judge missed only two of the twelve predictions by more than one point, in other words, in twelve predictions of $p = 1/5$, our judge scored six perfect hits, four near-hits, and two misses.

The therapist who had had thoroughgoing supervisory contacts with many of the therapists for whom this measure is available could not, however, identify the patterns, in terms of "liking" and "demanding". This might mean that the interpretation given to the patterns is not accurate, but it might also mean, that attitudes such as those tapped in this investigation, do not emerge in supervisory contacts. The above mentioned supervisor suggested the interpretation for UAS of "patient-centeredness or empathy of the therapist", that is, does he give or withold of himself. This is a possible interpretation which needs to be checked when more data can be collected.

Discussion

This paper has presented a method of quantification of certain aspects of countertransference for which we have suggested some interpretations. The sample of therapists is, however, too small, and uncontrolled to permit any definitive conclusions. Thus, for example, the point in treatment could not be equalized for all therapists, nor could provision be made that all sortings were done within the same time span or in the same order. Some patients saw the task as a psychological test, others were asked by their therapists to assist in this research. It is, therefore, all the more surprising that the results of the exploratory study show so much consistency. This might suggest that therapists' feeling reactions toward their patients are relatively stable over many treatment interviews, and perhaps from case to case. This has previously also been found in an investigation of wire recorded treatment sessions[1].

The significance which we can attach to the results of the study as well as to the method itself are:

1. That certain countertransference attitudes can be quantified and meaningfully interpreted, a step which might be found very useful in training and selection of therapists, and,
2. that countertransference attitudes are related to therapeutic competence.

While we cannot be certain about the correct interpretation of these phenomena, it seems fairly clear that they are important aspects of therapy. The main purpose of the present paper has been to demonstrate a relatively easily administered method for quantification of some apparently important therapist attitudes. This method may also be applied to quantification of similar phenomena, such as the transference relationship and the interrelationships between transference and countertransference. Such investigations are now in progress.

Summary

A method for quantification of certain countertransference attitudes has here been presented. The method involves the self-description of the patient by means of an array of statements which he is to sort in line with Q-technique methodology. The therapist is asked to predict the patient's self-description. He is also asked to

describe himself, and to describe how he would ideally like to be. The therapist's over- or underestimation of the similarity between his own and the patient's self-descriptions, and the estimated similarity between the patient's self-description and the therapist's "ideal" description is then used to estimate the nature and intensity of some countertransference feelings which have here been tentatively interpreted as falling on the dimension of empathy or acceptance-rejection and demandingness-supportiveness. An exploratory study is presented as an example of the method and the results which can be obtained.

References

1. Fiedler, F. E. A comparison of psychoanalytic, nondirective, and Adlerian therapeutic relationships. *J. consult. Psychol.*, 1950, *14*, 436-445.
2. Freud, S. The future prospects of psychoanalytic therapy. In *Collected Papers*, Vol. ii, London: Hogarth Press, 1924.
3. Murray, H. A. *Explorations in personality*. New York: Oxford Press, 1938.
4. Rogers, C. R. *Counseling and Psychotherapy*. Boston: Houghton Mifflin, 1942.
5. Stephenson, W. A. Statistical approach to typology: the study of trait universes. *J. clin. Psychol.*, 1950, *6*, 26-38.
6. Weiss, E. Manipulation of the transference relationship, in Alexander, F. and French T. *Psychoanalytic therapy*. New York: Ronald Press, 1946.

AN EVALUATION OF THE USE OF DIFFERENCE SCORES IN PREDICTION

J. R. WITTENBORN

Yale University

Introduction

In the Rorschach test and certain tests of mental deterioration a major emphasis is placed on the difference between two scores. This is done explicitly and formally so that for a sample of individuals there is available not only a set of scores which may be referred to as the y variable and another set of scores which may be referred to as the x variable, but also a third set of scores which may be referred to as the difference score variable and which comprises the difference between y and x for each individual in the sample. So far as the writer is aware, the logical and the empirical status of difference score variables has not been systematically investigated and the rational origins of the practice of using difference score variables are obscure.

The advantages peculiar to the use of difference scores have not been explicated by those who use such scores in clinical psychology, and psychologists differ greatly in the role they are willing to ascribe to them. In some studies it may be inferred that the differences between y and x provide a desirable "natural" weighting of y and x for predictive purposes. In most of the literature, however, discussions are handled in a way such as to imply that the difference between y and x, as a variable, possesses a predictive advantage which cannot be realized from a simultaneous use of the variables y and x. It is the purpose of the present discussion to examine this interesting possibility.

The Problem and a Proposed Solution

To what degree and under what conditions, if any, can the quantity $(y - x)$ be used as a variable to predict a portion of the variance of a criterion not predictable by weighting separately the variables y and x? It appears that a feasible approach to this question is possible by determining what portion of the variance of the variable $(y - x)$ is independent of y and of x. This approach employs two well known applications of correlation statistics:

A. If for the variable $(y - x)$ there is a reliable portion of variance independent of both y and x, this portion of $(y - x)$ variance could be used in predicting for some criterion a portion of its variance not predictable by an employment of y and an employment of x.

B. The variance of the difference score which is independent of both y and of x is considered to be equal to the quantity $(1 - R^2_{(y-x).yx})$.

These two applications may be integrated and summarized by the following equation:

$$S^2_{(y-x).yx} = (1 - R^2_{(y-x).yx}) \qquad (1)$$

where $S^2_{(y-x).yx}$ is the variance of $(y - x)$ which is independent of both y and x, and $R^2_{(y-x).yx}$ is the portion of the variance of $(y - x)$ which is predictable from y and x.

In order to apply this approach it is obviously necessary to know in any given instance what portion of the difference score variance is predictable from y and what portion is predictable from x. A convenient determination of these portions is possible if it is known to what degree y is correlated with the $(y - x)$ variable and to what degree x is correlated with the $(y - x)$ variable. These correlations, $r_{y(y-x)}$ and $r_{x(y-x)}$, may be computed directly after the differences between y and x have been determined. Computing the differences and the correlations is a tedious procedure

and identity equations were derived by the writer to provide a convenient, direct determination of the portion of the difference score variance which is predictable from y and from x. Assuming that the correlation between y and (y − x) is linear, we may write the following identity equation:

$$r_{y(y-x)} = \frac{\Sigma_{y(y-x)}}{nS_yS_{(y-x)}} \tag{2}$$

By means of a lengthy derivation the following new identity may be written:

$$r^2_{(y-x)y} = \frac{(S^2_y - S_yS_xr_{xy})^2}{S^2_y(S^2_y - 2S_yS_xr_{xy} + S^2_x)} \tag{3}$$

It is also possible to write a similar identity equation for $r^2_{(y-x)x}$. When S_y and S_x are considered to be equal to each other and equal to one, it is possible to write the identity equations in the following form:

$$r^2_{y(y-x)} = r^2_{x(y-x)} = \frac{(1 - r_{yx})}{2} \tag{4}$$

When this expression for $r^2_{y(y-x)}$ and $r^2_{x(y-x)}$ is substituted into the usual formula for R^2*, it is apparent that $R^2_{(y-x).yx}$ equals 1.00. It is concluded, therefore, that:

$$S^2_{(y-x).yx} = (1 - R^2_{(y-x)yx}) = 0.00 \tag{5}$$

The variable (y − x) is completely predictable from y and x and no portion of its variance is independent of y and of x. A difference score variable such as (y − x) cannot be used to predict a portion of the variance of a criterion which is not predictable by a use of variable y and variable x.

* $R^2_{(y-x).yx} = \dfrac{r^2_{x(y-x)} + r^2_{y(y-x)} - 2r_{y(y-x)}r_{x(y-x)}r_{yx}}{1 - r^2_{yx}}$

Discussion

An empirical verification of the determinations provided by equation 3 is given in Table 1. From equation 3, it may be seen that the portion of the variance of the

Table 1. Correlations computed directly from data for four samples.

Sample	A	B	C	D
N	100	100	100	37
$\bar{X}$	−.05	.00	−.03	23.22
$\bar{Y}$	.00	.00	.00	25.54
S_x	2.28	2.45	2.33	6.84
S^2_x	5.21	6.02	5.45	46.37
S_y	2.25	2.43	2.47	7.45
S^2_y	5.08	5.92	6.08	55.57
r_{xy}	.29	.65	.65	.50
$r_{x(y-x)}$	−.60	−.43	−.36	−.43
$r_{y(y-x)}$	.59	.41	.47	.56
Correlations determined by equation 3.				
$r^2_{x(x-y)}$	.36	.18	.13	.18
$r_{x(x-y)}$	±.60	±.43	±.36	±.43
$r^2_{y(y-x)}$	.35	.17	.22	.31
$r_{y(y-x)}$	±.59	±.41	±.47	±.56

(y – x) variable which may be predicted from y (or from x) is a function not only of the magnitude of the variance of y relative to the magnitude of the variance of x, but also of the degree to which y and x are correlated, r_{yx}. Table 2, which is based on an employment of Formula 3, shows the portion of the (y - x) difference score variance which may be predicted from y under a variety of conditions. When the variance of x is not equal to the variance of y (when $S^2_x \neq S^2_y$), the amount of the difference score variance which may be predicted from y is different from the amount of the difference score variance which may be predicted from x. Accordingly, Table 3 has been prepared to show the portion of the difference score variance which may be predicted from x under different conditions.[1]

TABLE 2. THE PORTION OF S^2 (y–x) PREDICTABLE FROM Y.

$\frac{S^2_y}{S^2_x}$	r = −1.00	r = −.90	r = −.70	r = −.50	r = .00	r = .50	r = .70	r = .90	r = 1.00
.00	——	——	——	——	——	——	——	——	——
.25	1.00	.91	.74	.57	.20	.00	.07	.46	1.00
.50	1.00	.93	.79	.66	.33	.05	.01	.16	1.00
.75	1.00	.94	.83	.71	.43	.15	.06	.01	1.00
1.00	1.00	.95	.92	.75	.50	.25	.15	.05	——

Each cell gives the value of $r^2_{(y-x)y}$ for the respective values of $\frac{S^2_y}{S^2_x}$ and r_{yx}

TABLE 3. THE PORTION OF $S^2_{(y-x)}$ PREDICTABLE FROM X.

$\frac{S^2_y}{S^2_x}$	r = −1.00	r = −.90	r = −.70	r = −.50	r = .00	r = .50	r = .70	r = .90	r = 1.00
.00	——	——	——	——	——	——	——	——	——
.25	1.00	.98	.93	.89	.80	.75	.77	.86	1.00
.50	1.00	.97	.90	.83	.67	.53	.50	.58	1.00
.75	1.00	.96	.87	.78	.57	.36	.29	.25	1.00
1.00	1.00	.95	.92	.75	.50	.25	.15	.05	——

Each cell give the value of $r^2_{(y-x)x}$ for the respective value of $\frac{S^2_y}{S^2_x}$ and r_{yx}

The results of the present analysis are interpreted to mean that the use of such quantities as M – C in the Rorschach represents no advantage which is not inherent in the direct use of M and C. The logic employed here is considered to apply equally well to other more elaborate types of difference score analyses.

When only two variables are involved, a difference score (although difficult to comprehend fully) may provide a kind of economy in the sense that one variable is substituted for two. This type of economy does not exist for more than two variables because for n variables there are $\frac{n(n-1)}{2}$ possible difference scores. It should also be noted that either y or x could have a high predictive value which would be obscured

[1]Respective values from Table 2 and 3 may be employed to show that $R^2_{(y-x).yx}$ is equal to one regardless of the magnitude of the variance of y relative to the variance of x and regardless of the degree to which y and x are correlated.

if y and x as variables were disregarded and predictions were based solely on the difference between them. In general the use of difference scores, patterns of differences, or profile analyses appears to be uneconomical for purposes of prediction.

The present discussion should not be taken as evidence that difference scores should never be used or that their use is invariably uneconomical. The primary purpose of the present discussion is to illuminate the fact that the use of difference scores confers no unique advantages. From the standpoint of economical use of data, difference scores are likely to prove disadvantageous because they in effect apply an arbitrarily weighted use of the two variables. Since an ordinary difference score disregards validity differences between the two original variables and fails to make any adjustment for differences between them with respect to mean and standard deviation, an investigator who uses difference scores profitably can ascribe his results to fortuitous circumstances and not primarily to the fact that he is using a difference score technique. The procedure which guarantees an optimal use of two variables for predictive purposes is the multiple regression procedure; this, as is well-known, involves a weighting of the two variables on the basis of their relative validity and on the basis of the relative magnitude of their units.

Summary and Conclusion

It has been our purpose to explore the possibility that the difference between two scores may be employed to predict a portion of the variance of a criterion which cannot be predicted from the simultaneous use of y and x as separate variables. It is indicated by the present analysis that a difference score has no predictive value superior to the predictive values of the initial scores. It is suggested, moreover, that the use of difference scores for predictive purposes possesses no unique merits, is uneconomical where more than two variables are involved, and in some situations may obscure the predictive merits of the variables. It is recommended that a weighting of two variables based upon a multiple regression procedure is in all cases equal if not superior to the use of difference scores per se.

CATINUA, CATEMENSIONS, CATESCALES

FRANK M. DU MAS

*University of Texas**

Introduction

This paper introduces several new concepts and indicates their utility for science in general and clinical and applied psychology in particular. This is the first of two papers on this subject. In this one we attempt a rational analysis of certain concepts and their associated phenomena. In the second paper we shall present the operational procedures for the immediate and practical application of the analysis presented in the present paper.

The significance of this paper for clinical and applied psychology lies in the fact that it now seems possible to use such information as is usually included in case histories, biographical data sheets, interview forms, etc., in two ways. (We shall use the neutral term "information files" to designate such sources of information.) The first way, and the method now in general use, is to have the psychologist become immersed in the information given and then make certain judgments or generalized evaluations of the individual. Such evaluations are of the very greatest importance and will probably remain as the basic considerata for evaluating people in quandaries. But because this is true it should not blind us to the other possible ways of using the *same information* in essentially the *same form*. Information files contain such a wealth of data that it seems desirable to wring every bit of significance out of them—even if we have to turn to quantitative or semi-quantitative methods.

Now the application of quantitative methods to such data is very difficult. The construction of planned information files which are susceptible to practical quantitative manipulation might easily destroy the most important aspect of them, namely, their rich description of the individual. We shall attempt to assist the evaluator—not obviate him—by indicating the fundamental nature of his data, and by indicating new methods of organizing, constructing, and evaluating information files.

Now information files, in the author's opinion, are composed of a rather chaotic aggregation of descriptive data that is at times categorical in nature (such as sex, religion, vocation, birthplace, etc.) and at other times dimensional in character (such as height, IQ, income, education, etc.). It would seem of great importance to systematize and order these data so that a more coherent view of the individual is obtained and at the same time quantify the sub-domains of information. We hope to lay out a guide for scientists who some day[1] may attempt to construct information files that have the characteristics which we have said to be desirable.

Definitions

For centuries philosophers and scientists have used the concepts, if not the terms, of category, continuum, dimension, and scale. The possible definitions of these terms are many and complex. After much thought and study it seemed that the definitions given below have the greatest relevance to the present discussion.

Category: a term that denotes the class character of phenomena which permits differentiation between phenomena in regard to kind rather than degree, quantity, or magnitude. Examples of categories are: cars, cabbages.

*The writer was on contract to the Office of Naval Research and this paper resulted from work done for ONR in research on scientific personnel. The writer wishes to thank Dr. Ralph M. Hogan for his help and encouragement. The ideas and opinions expressed are the author's and should not be construstrued as those of the Navy Department.

[1]The author has applied many of the concepts and procedures that are presented in this paper to (a) the evaluation of the proficiency of scientists and (b) a description of their work. Many of the examples given in this paper are, therefore, practical examples of how the present principles are applied.

Continuum: uninterrupted extent such that parts that are differentiable in space or in a spatial model exhibit the same class character. Examples of continua are: a piece of wire, a cloudless sky.

Dimension: a continuum exhibiting the same class character throughout but changes in regard to a serial order of degree, quantity or magnitude. Examples of dimensions are: length, weight.

Scale: an operational reference frame whereby reliable allocation of phenomena to a serial position within a dimension is facilitated. Examples of scales are: Centigrade, Fahrenheit, and Kelvin temperature scales.

The Catinuum

Consider now a long wire. Adjacent or non-adjacent parts all exhibit the same class character. But the psychological visual spectrum is different. Any two adjacent parts that are barely differentiable in space exhibit the same class character. There is also uninterrupted extent. But non-adjacent parts may not exhibit the same class character. The psychological visual spectrum has at least four modal class characters: red, yellow, green, and blue. We believe a hybrid concept, that is, a concept somewhere between category and continuum, would more adequately describe the psychological visual spectrum. We suggest the term *catinuum*, defined as follows: uninterrupted extent such that adjacent parts which are barely differentiable in space or in a spatial model exhibit the same class character but when parts are sufficiently non-adjacent they exhibit different class characters.

The Catemension

Consider the consecutive octaves making up the psychological musical scale. Musicians can easily allocate a note to an "octave class." The class character is rather sharply bounded. But an octave higher up the scale is consistently different from a lower octave. This unique octave "dimension"—not a pitch dimension, though it may vary with pitch—might be better referred to as a *catemension*, defined as follows: a dimension which is indicated only by a set of ordered categorical phenomena.

The Catescale

Now, it seems possible that the operation of allocation to a class can serve as measurement. Consider an obvious physical example. A judge might order a group of objects as given in Figure 1 in regard to their weights. These terms, which represent categories, designate positions along a scale. Another individual would obtain a primitive "measurement" of the weight of each object from a conglomeration of sofas, houses, carrots, axes, and cars by the operation of allocation to its appropriate category. See Figure 1.

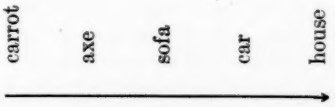
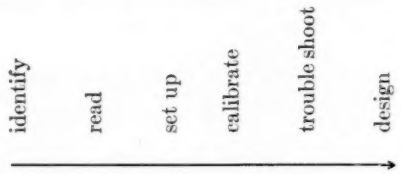


FIGURE 1. Physical example of a catescale.

Consider now a psychological example. Suppose that a psychological scientist makes observations of physical scientists using a particular instrument—say, an oscilloscope. For several days he notes all of the activities of many scientists using this instrument in their work. Suppose further that the range of proficiency in the

use of this instrument as exhibited by these scientists is maximal. After a thorough analysis of his data he finds that the behavioral categories given in Figure 2 are associated with the degree of proficiency with the oscilloscope. See Figure 2.

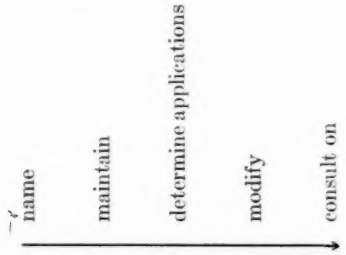


Dimension: Proficiency with oscilloscope.

FIGURE 2. Psychological example of a catescale.

Now then: a judge (the allocator) need merely to assign an individual evaluated (the allocatee) to those categories which describe the allocatee. Or it is possible that an allocatee placed in a particular category is also generally placed in all categories below it. If this is true, then the highest category to which an individual is assigned is an indication of his degree of proficiency with the oscilloscope.

We have noted that there may be more than one scale used to measure the same underlying dimension. This is also true of the catescale. All catescales are catinua but not all catinua are catescales. See Figure 3.



Dimension: Proficiency with oscilloscope.

FIGURE 3. Alternate catescale to Figure 2.

We now suggest the third hybrid concept, the *catescale*, defined as follows: a set of categories which are serially and uniquely ordered by reference to the degree, quantity, or magnitude exhibited by each category of a referent dimension.

CATESCALE VS. RATING SCALE

The catescale and the rating scale should not be confused. They differ in many ways. Some of the more important differences are given below.

a. A rating scale indicates a generalized estimate of a ratee by a judge; the catescale is non-commital and merely indicates the class of behavior which the allocator has observed the allocatee to exhibit. That is, the rating scale yields an estimate—the catescale, an observation.

b. A rating scale is a verbal statement of degree; a catescale is a set of categories that say nothing directly about degree.

c. It is difficult to see how very different rating scales of the same underlying dimension are possible; one can have quite different catescales of the same referent dimension. See Figures 2 and 3.

d. Because of certain perceptual characteristics of human beings the catescale has certain other advantages over the rating scale. When the categories are unordered in the catescale, the allocator may not have insight into what is being measured and this factor would seem to objectify the measurements made from catescales in that it should decrease the halo effect. See Table 1.

TABLE 1. COLUMN A CONTAINS THE BEHAVIORAL CATEORIES OF FIGURE 2 ORDERED IN THE FORM OF A CATESCALE. COLUMN B CONTAINS THE SAME CATEGORIES UNORDERED.

A	B	
Identify	Design	
Read	Read	
Set up	Calibrate	
Calibrate	Set up	: the oscilloscope
Trouble shoot	Identify	
Design	Trouble shoot	

Again, we might modify slightly and imbed the unordered behavioral categories in a mass of irrelevant categories. Such a check list would be scored for the relevant behavioral categories only. It is probable that the allocator might not discern the underlying dimension in such a conglomeration. See Table 2.

CATESCALES VS. TESTS

The catescale has certain advantages over the psychological test. Some of the more important advantages follow.

a. Where cost is important factor, the catescale is far superior to the test.

b. Where time is important, the catescale is far superior to the test.

c. Practical application of the catescale to a large number of variables is easy; the same number of tests to measure the same variables would often be impossible of application.

SCALE CHARACTERISTICS OF THE CATESCALE

We are concerned here with the question of procedures for the construction of a catescale. Our goal is to recommend procedures which would seem to yield the following results:

a. A unique sequence for the categories making up the catescale.

b. The scale separation between categories.

c. If possible, give a general quantitative meaning to the catescale.

Where there exists an apparent phenomenal order among the categories we may apply the standard psychological scaling procedures such as: method of rank order, rating scale methods and the method of paired comparisons. There arises, however, a significant problem when we attempt to scale categories which exhibit *no* apparent phenomenal order. The author has worked out a technique which may be used for either type of category.

TABLE 2. THE CATESCALE OF FIGURE 2 MODIFIED SLIGHTLY, UNORDERED AND EMBEDDED IN A MASS OF IRRELEVANT CATEGORIES.

Can design	
Know cost of	
Makes readings of	
Likes to work with	
Calibrates efficiently	: the oscilloscope
Talks about frequently	
Can set up	
Is careful with	
Trouble shoots	
Can identify	

Reliability and Validity of Catescales

It seems that the techniques already in existence are applicable in determining the reliability and validity of catescales. Some special problems may arise but, in our opinion, they will not be insurmountable.

Parallel Catescales

Parallel catescales for the same dimension might be obtained in several ways. The method suggested below allows parallel forms to be constructed without additional work, once the scale values are obtained. Let us assume that the method of rank order was used as the psychological scaling procedure for the categories in Table 3. First, a category is chosen at each end of the catescale for each of the parallel catescales.

Table 3. Construction of parallel catescales from a single original catescale. Hypothetical rank order scale values are used.

Scale value	*Category*	*Scale value*	*Category*
		1.4	Name
1.4	Name	2.8	Read
1.7	Identify	3.3	Set up
2.8	Read	5.9	Calibrate
3.0	Maintain	7.9	Modify
3.3	Set up	10.4	Design
5.6	Determine application		
5.9	Calibrate		
7.6	Trouble shoot	1.7	Identify
7.9	Modify	3.0	Maintain
10.4	Design	5.6	Determine application
10.4	Consult on	7.6	Trouble shoot
		10.4	Consult on

This is necessary in order to make sure the ranges are approximately equal. Then categories are selected from alternate positions in the original catescale and placed in their same relative positions in the parallel catescales. See Table 3.

A Special Case

Let us now consider the scaling of categories which do not exhibit an apparent phenomenal order. See Table 4.

Table 4. Categories that might be scaled where no apparent phenomenal order exists. Such data are usually included in information files.

	Ohio (birthplace)
	Carpenter (vocation)
	Male (sex)
Categories:	Catholic (religion
	White (race)
	Married (marital status)
Closed interval variables considered as categories:	Parents divorced (home life)
	IQ: 120-140
Open interval variables considered as categories:	Age: 25-35
	Height: 6 feet or over
	College Graduate

It is interesting to note that even when a phenomenal order does exist among the categories, these categories may be manipulated so that no apparent phenomenal order exists or whose order is very difficult to perceive. See Table 2.

Application

Theoretically, it seems that the catescale technique is applicable wherever rating scales or tests are applicable. There are two instances where the catescale might well permit new and radical solutions to old, tough problems.

First, the critical evaluation of malingering and falsification. This problem has plagued clinical psychologists recently and psychiatrists, physicians, insurance companies and administrators for a long time. Present methods of detecting malingering in regard to the sensorium are rather adequate. They are very inadequate, however, in cases of backache, headache, visceral pains and also in determining the genuineness of fears, dissociated states, psychotic episodes, compulsive behavior, etc. The evaluation of the last three types of behavior is relevant to legal questions concerning the state of an individual at the time he committed a crime. Because categories can be scaled where no apparent phenomenal order exists, catescales should be most difficult to falsify.

Second, the catescale technique suggests a scientific and systematic method for constructing certain types of data-gathering instruments. The author knows of no major theoretical contribution to the construction of case histories, information files, interview forms, etc. The catescale technique indicates a new procedure for the construction, organization, and evaluation of case histories, biographical data sheets, interview forms, questionnaires, job analyses, job descriptions and other analytic, multivariate data-gathering instruments of psychosocial phenomena.

A SCALE OF NEUROTICISM: AN ADAPTATION OF THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY[1]

JOHN F. WINNE, PH.D.

V. A. Neuropsychiatric Hospital
Northampton, Mass.

Introduction

One of the most recent, and among the best of personality inventories is the Minnesota Multiphasic Personality Inventory (1, 2, 3, 4) which purports to measure many aspects of personality by scoring various combinations of items. Although this scale has found great use in its present form in clinics and hospitals, it was thought desirable to determine if there were items which would distinguish between normals, on the one hand, and a group of neurotics of mixed diagnoses. Should such a set of items be found, a two-fold need could be relieved. First, a short scale useful in screening out all types of neurotics could be devised by combining the items in some fashion. Secondly, these items could form the basis for a comparative factor analytic study of personality organization in normals and neurotics.

[1]The study reported in this paper is an extension of part of an experiment performed in partial fulfillment of the requirements for the degree of Doctor in Philosophy at the University of Pennsylvania under the direction of Professors Malcolm G. Preston, Robert A. Brotemarkle, and Morris S. Viteles. Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are a result of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

Selection of Subjects

The 560 subjects whose records were used in this study were male white veterans who were given the MMPI as part of a program of vocational advisement or in the course of the psychological screening examination at the Philadelphia VA Mental Hygiene Clinic.[2] These subjects were divided into four groups of 140 men each; two such groups were composed of neurotics, while the other two groups were normal with respect to psychiatric illness. One pair of normal-neurotic groups was used in the validation process; the other set was used in the cross-validation. Subjects were classified as neurotic if they carried a pensionable rating by the VA of such disability or if they were so diagnosed by their attending psychiatrist. Subjects were classified as normal if there was no record in any of the available VA files of their having requested or received any psychiatric treatment before, during, or since military service.

The subgroups of normals and neurotics were matched for certain items of biographical data at the time they first took the MMPI. These variables include intelligence, age, education, occupation, and marital status. Tests of Chi-square and critical ratio indicate that all groups are homogeneous with respect to these variables. The typical subject has average intelligence, is about 25 years of age, has completed the tenth grade of school, and has about an equal chance of being married or single. Most of the subjects were employed in an on-the-job training program for skilled trades, or were students in high school and college. For details of the matching, and the distribution statistics, the reader is referred elsewhere[6].

Development of a Scale of Neuroticism

The 117 items which make up the Hysteria, Depression, and Hypochondriasis scales of the MMPI were selected for this analysis because it was felt that they should contribute a large number of items which would distinguish between normals and neurotics. The responses made by each subject to each of the items was recorded and the distributions of True, False, and Cannot Say responses to each item were ascertained for one group of 140 normals and the comparable group of 140 neurotics. Two sets of comparison by Chi-square were made for each item to find the differentiating ability of that item: True vs. False & Cannot Say; True & Cannot Say vs. False.

Only 33 items were found which distinguished between normals and neurotics at the .01 level of confidence. The 30 items with highest Chi-square values were then combined into a Scale of Neuroticism, presented in Table 1 (with items arranged in order of their differentiating ability), showing the number of the item in the MMPI booklet, the deviant response, the percent of each group giving the deviant response, and the Chi-square value attaching to the deviant response.

It will be noted that Item 208 was more commonly answered False by neurotics, while McKinley and Hathaway report it as a *normally false* item. For all other items, the direction of the original standardization was verified. Also to be noted is the preponderance of physical symptoms, which apparently arise from two sources: in the first place, the three scales of the MMPI which were used as the universe of symptoms for this study are heavily weighted with such symptomatology. Secondly, we are dealing here with a population of veterans who are receiving monthly pensions for their neuroses, and it is to be expected that physical complaints would be more readily acceptable to these men than complaints of emotional dysfunctions.

---

[2]The writer wishes to acknowledge the assistance of the following persons who made this experiment possible by releasing records and test results: John E. Davis, M.D., formerly Chief Psychiatrist, and Charles M. Morris, Ph.D., formerly Chief Clinical Psychologist of the VA Mental Hygiene Clinic, Philadelphia; Mr. Walter E. MacDonald, Chief, Advisement and Guidance Section, VA, Philadelphia; Mr. Daniel N. Wiener, Chief, Advisement and Guidance Section, VA, Minneapolis; and Mr. Seth Biemiller, Fleisher School, Philadelphia.

TABLE 1. A SCALE OF NEUROTICISM

Rank Order	*Booklet Number*		*Deviant Response*	*Percent Normal*	*Deviant Neurotic*	*Chi-square (d.f. = 1)*
1	43	My sleep is fitful and disturbed.	T	21.43%	47.86%	21.5943
2	159	I cannot understand what I read as well as I used to.	T	19.29	44.29	20.1777
3	44	Much of the time my head seems to hurt all over.	T	14.29	37.14	19.1453
4	3	I wake up fresh and rested most mornings.	F	47.14	71.43	17.1042
5	72	I am troubled by discomfort in the pit of my stomach every few days or oftener.	T	21.43	43.57	16.6451
6	186	I frequently notice my hand shakes when I try to do something.	T	37.86	61.43	16.5151
7	189	I feel weak all over much of the time.	T	18.57	40.71	16.4565
8	9	I am about as able to work as I ever was.	F	40.00	63.57	15.5770
9	236	I brood a great deal.	T	25.00	45.71	15.0165
10	107	I am happy most of the time.	F	25.71	47.86	14.7595
11	103	I have little or no trouble with my muscles twitching or jumping.	F	29.29	51.43	14.2589
12	41	I have had periods of days, weeks, or months when I couldn't take care of things because I couldn't "get going".	T	30.71	52.86	14.1094
13	242	I believe I am no more nervous than most people.	F	44.29	66.43	13.8880
14	29	I am bothered by acid stomach several times a week.	T	18.57	35.71	13.7200
15	2	I have a good appetite.	F	16.43	35.71	13.5080
16	190	I have very few headaches.	F	28.57	50.00	13.4759
17	114	Often I feel as if there were a tight band about my head.	T	14.29	32.86	13.4013
18	175	I seldom or never have dizzy spells.	F	31.43	52.14	12.3475
19	47	Once a week or oftener I suddenly feel hot all over, without apparent cause.	T	12.86	30.00	12.2182
20	238	I have periods of such great restlessness that I cannot sit long in a chair.	T	54.29	74.29	12.1956
21	208	I like to flirt.	F*	41.43	61.24	12.0926
22	46	My judgment is better than it ever was.	F	29.29	49.29	11.7390
23	108	There seems to be a fullness in my head or nose most of the time.	T	23.57	42.14	11.0021
24	51	I am in just as good physical health as most of my friends.	F	34.29	52.86	9.8195
25.5	5	I am easily awakened by noise.	T	42.14	60.71	9.6650
25.5	191	Sometimes, when embarrassed, I break out in a sweat which annoys me greatly.	T	39.29	57.86	9.6650
27	178	My memory seems to be all right.	F	16.43	32.14	9.4007
28	68	I hardly ever feel pain in the back of the neck.	F	27.14	44.29	8.9600
29	76	Most of the time I feel blue.	T	23.57	39.29	8.6121
30	263	I sweat very easily even on cool days.	T	40.71	57.86	8.2303

*This result is in the opposite direction from that obtained in the standardization of the MMPI. See text.

When individual records were scored for deviant responses to the Scale, the number of such responses for neurotics in the validation group ranged from 3 to 30 (mean = 14.8; S.D. = 6.7) while the scores for the comparable normals ranged from 0 to 24 (Mean = 8.6; S.D. = 5.9). In the cross-validation groups, the range for neurotics was 1 to 30 (Mean = 14.1; S.D. = 5.6) and the range for normals 0 to 24 (Mean = 8.9; S.D. = 5.4). The biserial correlation between class membership and number of deviant responses was .554 for the validation group and .530 for the cross-validation group. The reliability coefficient for all normals, using product-moment correlation between random halves of the scale was .728, corrected by the Spearman-Brown formula to .843. The corresponding correlation for all neurotics was .757, corrected to .862. All these correlations are well above the .01 level of confidence.

A cutoff score was established at that point in the distribution of the validation group where the number of misidentifications reached a minimum. At this point, 11 or more deviant responses, 68.58% of the validation group and 66.79% of the cross-validation group were correctly identified. The records of a third group of 107 subjects, not otherwise used in this experiment, and not matched for background data, was also tested with this scale, using the same cutting score. Somewhat better results were obtained with this group; 71% were correctly identified. The misidentifications for all groups were fairly evenly divided between false negatives (normals called neurotic) and false positives (neurotics called normal) as is shown by Table 2.

TABLE 2. EFFECTIVENESS OF DIFFERENTIATION

	Validation		*Cross-Validation*		*Extras*		*Total*	
	N	%	*N*	%	*N*	%	*N*	%
False Positive	40	28.56	48	34.28	22	31.42	110	31.42
False Negative	48	34.28	45	32.14	9	24.32	102	32.17
Total Wrongs	88	31.42	93	33.21	31	28.97	212	31.78
Total Rights	192	68.58	187	66.79	76	71.03	455	68.22
Total	280	100.00	280	100.00	107	100.00	667	100.00

DISCUSSION

An examination was made of the biographical and diagnostic data for the false positives and false negatives in an attempt to find a rationale for these misidentifcations. Explanations for the false positives (neurotic called normal) are much more readily possible than for the false negatives, but it must be made explicit that these explanations are only tentative and conjectural in nature. It is well known and readily accepted that diffuse anxiety is reduced, in the neurotic, by the development of somatic symptomatology, either physiogenic or psychogenic in nature. May (5) expresses the idea as follows: "But when the conscious struggle can no longer be tolerated, either because of its increasing severity or because of its lack of success, symptomatic changes in the organism take place. These relieve the strain of the conflict and make a quasi- or pseudo-adjustment possible when the conflict cannot actually be solved. Thus it may be said that the symptoms are often ways of containing the anxiety; they are the anxiety in structuralized form" (p. 84). We should expect, then, to find many subjects with conversion symptoms in the false positive group since these patients would lack the diffuse or free floating anxiety measured by this adaptation of the MMPI. Examination of the false positives shows that of 110 subjects in the group, there are 13 stutterers, 12 cases of conversion hysteria, and 4 cases of psychogenic illness (ulcers, somnambulism, and migraine). Closely related to these, and explainable by the same reasoning, are 14 cases of mixed neurosis and 10 cases of unclassified neurosis. Apparently these persons did not have sufficient anxiety to be screened by the Neurotic Scale. A third related group, which

lacks sufficient anxiety consists of 20 mild anxiety states. Thus, we can account for 73 of 110 false positives. May further points out that when neurotic patients become ill in ostensibly physiogenic ways, anxiety also tends to decrease. Six of the remaining false positives have more or less physiogenic illnesses such as diabetes, near blindness, hammer toes, etc. These seem to serve the function of focusing the neurotic anxiety and thus eliminating diffuse symptomatology. We are still left, however, with 31 cases for whom no explanation can be found—a residual error of about 9%.

Turning now to the 102 false negatives (normals called neurotic) we have somewhat more difficulty in finding a rationale for misclassification. Misunderstanding seems to have played some part, since we find here 4 persons with intelligence test scores two or more standard deviations below the experimental mean. There is considerable doubt about the degree of normality in several cases also. Three cases were assumed to be normal only because all information was unavailable; they are all classified as neurotic by the present scale. Five cases have illnesses of a psychogenic nature (ulcers, asthma, low back pains, headache), while 5 more cases have hearing difficulty, giving support to the oft-repeated statement that deaf persons are prone to have personality maladjustment. Finally 13 subjects had personal problems of such degree that they voluntarily reported to a mental hygiene clinic where they were neither diagnosed as neurotic nor accepted for treatment. Thus we can account, to some degree, for 26 more of our false negatives. A third factor which seems to be involved in the misidentification of normals is the amount of normal anxiety present. We have seen that physical disease reduces neurotic anxiety in the neurotic, but, conversely, it increases normal anxiety in the normal. Support for this hypothesis is found in the fact that we find 32 normals with physical disability misclassified as neurotic: 5 malaria, 4 rheumatic fever, 2 loss of vision, and the remainder having a variety of physiogenic conditions. This leaves us with 40 cases for whom we cannot account, a residual error of about 12%. This is, in actuality, a small number when compared with the commonly accepted figure that from 20 to 30% of the general population are neurotic. Additional information as to cultural background and presence of physiogenic symptomatology might well reduce this residual error to a minimum.

We see, therefore, that although initial screening of known cases with this 30 item scale results in 32% error, further examination reduces the error to about 10%, suggesting that two sources of error must be controlled in any practical application. Information must be obtained as to the presence of psychogenic and physiogenic symptomatology which will aid us in determining the disposition of borderline cases. Secondly, it may be useful to create a borderline group, by setting upper and lower cutting points, which can be more intensively examined for psychiatric disability. Finally—a general limitation of all inventories—local norms should be obtained for each installation using the scale to insure that the results here reported are not merely artifacts of a specific environment.

Summary

1. Of the 117 items in the Hysteria, Depression, and Hypochondriasis scales of the Minnesota Multiphasic Personality Inventory, only 33 items were found to distinguish at the .01 level of confidence between a group of 140 male white neurotic veterans and a group of 140 normal veterans equated in terms of race, sex, intelligence, age, education, occupation, and marital status.

2. When the 30 items which distinguish at the highest level of confidence are combined into a scale with unit weights, we obtain a Scale of Neuroticism which correctly identified about two-thirds of known normals and known neurotics. The biserial correlation between number of neurotic responses and class membership is .554 for the validation group and .530 for a cross-validation of 280 similar veterans.

3. The reliability of the scale is .843 (split half, corrected) for the neurotic group, and .862, (split half, corrected) for the normal group.

4. The scale is less effective for screening out known neurotics with physiogenic and conversion symptoms or with mild and transient anxiety than it is for identifying more severe anxiety neurotics.

5. Normals with severe physical disorders, very low intelligence, deafness, or possible psychosomatic complaints (ulcers, asthma, etc.) are less apt to be correctly identified than those normals lacking these symptoms.

References

1. Hathaway, S. L. and McKinley, J. C. A multiphasic personality schedule (Minnesota): I. Construction of the inventory. *J. Psychol.*, 1940, 10, 249-254.
2. Hathaway, S. L. and McKinley, J. C. A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. *J. Psychol.*, 1942, 14, 73-84.
3. Hathaway, S. L. and McKinley, J. C. The Minnesota personality inventory: V. Hysteria, hypomania, and psychopathic deviate. *J. Appl. Psychol.*, 1944, 28, 153-174.
4. McKinley, J. C. and Hathaway, S. R. A multiphasic personality schedule: II. A differential study of hypochondriasis. *J. Psychol.*, 1940, 10, 225-268.
5. May, Rollo. *The Meaning of Anxiety.* New York: The Ronald Press. 1950.
6. Winne, John F. The factorial composition of normal and neurotic responses to an adaptation of "The Minnesota Multiphasic Personality Inventory". Unpublished Ph.D. Dissertation. On file, University of Pennsylvania Graduate School Office.

BRIEF TESTS OF INTELLIGENCE IN THE PSYCHIATRIC CLINIC

JOHN R. KNOTT, RICHARD G. CANNICOTT,
JOHN P. UMBERGER AND INA McD. BILODEAU

Division of Psychology
Iowa State Psychopathic Hospital, University of Iowa,
and Department of Psychiatry, State University of Iowa College of Medicine

INTRODUCTION

During the recent war many psychologists were faced with the problem of obtaining some kind of individual psychological impression of large numbers of men with only a brief interval for examination allowed per man. There is a continuing interest in such short tests, which indicates that they may have a useful place in the clinical psychologist's armamentarium. Under the impetus of W. A. Hunt [10, 11, 12], the Kent EGY and other tests from the Kent Battery of Emergency Tests[17 to 22] were used rather widely in the Navy. Aside from this military application of parts of the Kent Battery, it has not been extensively employed, except for one or two of the subtests. It appeared to the authors, from military use of the battery, that it might have merit in civilian clinical practice, where there is often need of quantitatively derived estimates of intelligence which are independent of the longer standard tests, and which will not invalidate their later administration.

The particular advantages of the Kent Battery are that they have been standardized, can be administered and scored in a short time, tap more than one area of psychologic functioning, are applicable at least in part to limited types of psychiatric material and there is some expectation of diagnostic variation, and have no items in common with other tests.

The present investigation was directly concerned with an evaluation of the usefulness of the Kent Battery in various subtest combinations in the psychiatric clinic. The criterion of usefulness was the predictive value of the short battery in the estimation of the intelligence level as defined by the Wechsler-Bellevue scale.

SUBJECTS

A sample of 100 patients was investigated. Eighty-two patients from the Out-Patient Clinic of the Iowa Psychopathic Hospital (79 of whom were subsequently admitted as in-patients), and an additional group of 18 patients from the obstetrical wards of the State University Hospital whose infants were to be placed in adoption homes were tested. The chronological ages of the group ranged from 13 to 64 years, with a median of 27 years; the distribution is moderately skewed to the right. The educational attainment ranged from Grade IV to graduate college status, with a median of 11 grades completed. The diagnostic groups (excluding the 18 obstetrical patients) of the psychiatric sample were: Schizophrenia, 10; InvolutionalPsychosis, 8; Manic Depressive Psychosis, 18 (of whom 15 were depressed); Psychoneurosis, 19; Psychopathic Personality, 17; Others, 10.

PROCEDURE

The following Kent subtests were used in various combinations:

Kent EGY (10 questions), a test of general information and reasoning.

Graduated Opposites A, a test involving vocabulary and maintenance of associative set.

Similarities, essentially a test of vocabulary.

Essential Difference, which presumably involves low level abstraction and maintenance of set.

Essential Property, presumably a test of higher level abstraction ability.

Directions B, which involves shifting of set, rapidity of associations and executions of tasks.

Arithmetical Reasoning, a test of ability to manipulate symbols within a learned matrix.

The first two tests were administered orally and recorded verbatim; all others were administered as individual written tests in the order given above. The average time required for administration of the seven tests was 25 minutes, although the median time fell between 15 and 20 minutes; the difference may be attributed primarily to depressed patients who greatly extend the average time.

The battery of short tests was given, for the most part, prior to administration of the Wechsler-Bellevue Intelligence Scale, which was the criterion instrument in this study. The interval between tests ranged from zero to 98 days, the median interval being 3.25 days. All tests were administered by staff members of the Division of Psychology of the Psychopathic Hospital.

Each subtest of the brief battery was scored on the basis of Kent's published norms[21]. These gave *age scores* which, to avoid the implication of mental age, have here been termed *derived scores*. Derived scores were available for 2-minute timed tests on *Arithmetical Reasoning* and *Directions B*. *Similarities*, *Essential Difference*, and *Essential Property* were scored on the basis of the available untimed norms.

From the seven subtest scores a "mean derived score" was computed for each patient, and an "intelligence level" was assigned. After some preliminary study it appeared that a mean derived score of 12 on the Kent Battery was equivalent to an *average* intelligence level, as defined by the Wechsler-Bellevue Scale, and that one point variations below and above this (i.e. 11 and 13) were apparently equivalent to Dull Normal and Bright Normal. One point variations below and above these values (i. e. 10 and 14) consequently would be Borderline and Superior, while scores of 9 or below would be Mentally Defective.

Results and Discussion

1. When the mean derived scores of the seven-test Kent Battery were correlated with the scores derived from the Wechsler-Bellevue Scale, a correlation coefficient of .872 was obtained (Table 1). The correlations with the Wechsler Verbal and Performance scales were .865 and .798 respectively. When two tests, *Similarities* and *Essential Difference*, were dropped to form a five test battery, the correlation was .864 with the Wechsler Full Scale. When the following four subtests, Kent *EGY*, *Essential Difference*, *Directions B* and *Arithmetical Reasoning*, were combined into a battery the correlation was .848 with the Wechsler Full Scale, and .832 and .789, respectively, with the Verbal and Performance Scales. All these correlations (Table 1) are significant beyond the one per cent level of confidence.

Table 1. Correlations Between Kent Emergency Battery and Wechsler-Bellevue Intelligence Scale (N = 100).

Kent Battery	*Correlation*	*Standard Error*	*1% Confidence Interval*
	Wechsler-Bellevue Full Scale		
7-test means	0.872	9.86	.75 – .91
5-test means*	0.864	10.06	.78 – .92
4-test means**	0.848	10.67	.75 – .91
	Wechsler-Bellevue Verbal Scale		
7-test means	0.865	9.76	.78 – .92
4-test means**	0.832	10.83	.73 – .90
	Wechsler-Bellevue Performance Scale		
7-test means	0.798	11.30	.68 – .88
4-test means**	0.789	11.49	.67 – .87

*Includes EGY, Grad. Opposites, Essential Properties, Directions B, and Arithmetical reasoning.
**Includes EGY, Essential Differences, Directions B, and Arithmetical Reasoning.

2. Of the 76 patients classified as average or above on the Wechsler-Bellevue, only 6 were incorrectly designated as below average intelligence by the short battery of tests. On the other hand, of the 24 patients classified as below average intelligence by the Wechsler, only 2 were designated as average by the short battery, none as above average. This would appear to provide sufficiently accurate discrimination for use in the psychiatric clinic. Obviously, in practice, one retests with a standard instrument if the short scale suggests a level below that expected, exactly as one should retest if a "short Wechsler"[27] were to fall below a certain specified value.

3. The inter-test correlations, between the subtests of the Kent battery, are presented in Table 2. They indicate moderate correlation between most of the subtests, but in some instances the relationship is lower. On *a priori* grounds, it appears that a variety of functions is being tapped by this battery, although a common factor is definitely present.

TABLE 2. KENT BATTERY SUBTEST INTERCORRELATIONS (N = 100).

	EGY	Grad. Opp.	Simil.	Esst. Diff.	Esst. Prop.	Direc. B.	Arith. Reas.
E G Y		.648	.646	.577	.593	.501	.714
Graduated Opp.	.648		.717	.541	.590	.628	.721
Similarities	.646	.717		.690	.612	.627	.650
Essential Diff.	.577	.541	.690	.	.587	.512	.516
Essential Prop.	.593	.590	.690	.587		.559	.651
Directions B	.501	.628	.627	.512	.559		.661
Arithmet. Reason.	.714	.721	.650	.516	.651	.661	

4. The correlations between the subtests of the Kent battery and those of the Wechsler-Bellevue are shown in Table 3. The moderate relationship indicated among these various subtests is similar, in degree, to that shown in the Kent Battery above, as well as to the inter-test correlations which Wechsler reported for his intelligence scale[32]. In this respect, the two scales appear to tap similar areas of functions.

TABLE 3. KENT BATTERY AND WECHSLER-BELLEVUE SCALE SUBTEST INTERCORRELATIONS (N = 95 IF NOT OTHERWISE INDICATED).

Wechsler Bellevue	Kent Emergency Battery						
	EGY	Grad. Opp.	Simil.	Ess. Diff.	Ess. Prop.	Dir. B	Arith. Reas.
Information	.735	.637	.698	.546	.635	.571	.736
Comprehen.	.692	.578	.627	.500	.624	.472	.660
Similarities	.658	.611	.701	.578	.644	.515	.645
Digits	.596	.637	.609	.543	.595	.492	.617
Arithmetic N = 94	.681	.618	.641	.540	.672	.518	.736
Vocabulary N = 58	.685	.595	.676	.494	.623	.586	.714
Pic. Arr.	.520	.507	.572	.488	.535	.434	.489
Pic. Comp.	.659	.599	.517	.492	.542	.483	.582
Block Design	.700	.565	.646	.541	.597	.554	.669
Obj. Assem.	.480	.413	.481	.413	.465	.359	.502
Digit Sym.	.409	.423	.476	.455	.461	.459	.439

Summary and Conclusions

A battery of seven subtests, apparently measuring a related variety of functions including general information, reasoning, vocabulary, associative set, flexibility of set, and abstract symbolic capacity, drawn from the Kent Battery of Emergency Tests allows adequate prediction of intelligence level for use in preliminary psychiatric evaluation and decision. This is accomplished at a saving in time in comparison with, and without invalidating parts of, standard intelligence tests. As few as four of the subtests may be used.

References

1. Hunt, W. A., *et al.* The clinical possibilities of an abbreviated intelligence test. *J. consult. Psychol.*, 1948, 12, 171-173.
2. Hunt, W. A., *et al.* The validity of some abbreviated individual intelligence scales. *J. consult. Psychol.*, 1948, 12, 48-52.
3. Kent, G. H. Oral test for emergency use in clinics. *Ment. Meas. Monogr.*, 1932, 9, 0. 50.
4. Kent, G. H. Written tests for the clinic. *J. genet. Psychol.*, 1934, 44, 49-68.
5. Kent, G. H. Self-derived norms for institutions. *Psychol. Record*, 1939, 3, 195-207.
6. Kent, G. H. Emergency battery of one minute tests. *J. Psychol.*, 1942, 13, 141-164.
7. Kent, G. H. Tentative norms for emergency battery. *J. Psychol.*, 1943, 15, 17-149.
8. Kent, G. H. Additional norms for emergency battery. *J. genet. Psychol.*, 1945, 67, 17-26.
9. Rabin, A. I. A short form of the Wechsler-Bellevue test. *J. appl. Psychol.*, 1943, 27, 320.
10. Wechsler, D. *The Measurement of Adult Intelligence.* Baltimore: Williams & Wilkins, 1944.

THE IMMEDIATE TEST*

RAYMOND CORSINI

Wisconsin Dept. of Public Welfare
Division of Correction

It is often useful in clinical work to obtain a quick objective estimate of intelligence. For this purpose a number of short individual scales have been constructed *de novo* or adapted from longer intelligence tests. Use of these instruments is generally deplored by the very clinicians who employ them, but the realities of some clinical situations make the employment of short scales necessary. Further attempts to create suitable short individual intelligence tests are worth-while in view of their usefulness and the shortcomings of present tests. To the end of attempting to create a more suitable test, the writer established a list of specifications for an ideal short test, to be called the Immediate Test. As the name implies, this test should immediately, economically, and conveniently assist in the determination of relative degrees of intelligence.

Specifications

Time. The single most important non-measuring element of an Immediate Test is conservation of the examiner's time. Five minutes should be the maximum average examination time from the moment of decision to use the test to the final score.

Convenience. An Immediate Test should require no special preparation and a bare minimum of material.

Applicability. An Immediate Test should be applicable for a large number of people. It should not unduly handicap individuals who have personality, sensory, or motor deviations.

Flexibility. An immediate test should permit even further condensation for those deplorable situations when an ultra-immediate judgment is desired.

Testability. An Immediate Test should create an immediate favorable attitude towards the test situation. The subject should understand what to do in a minimum time and be able to respond effectively without any difficulty.

Efficiency. An Immediate Test must be efficient. Very little time should be wasted either by the subject or the examiner. A large number of measuring points should be provided. The individual should be examined only within his effective range.

Accuracy. The Immediate Test should possess satisfactory validity and reliability.

Description of the Immediate Test

The Immediate Test takes 3½ minutes per subject from the moment of beginning to the final IQ score. Over 90% of subjects are tested and rated within five minutes. For the experienced examiner only a single printed sheet of paper and a pencil are needed for the complete evaluation.

The test is applicable to anyone for whom a standard individual test such as the Wechsler or the Binet is suitable. It can be condensed to 2/3, 1/2, 1/3 or 1/6th its normal length. Most subjects like to take this test and begin without any hesitation since it is a familiar and interesting situation. Scoring is done along with the administration. The number of measuring points are almost as many as for some tests ten times as long.

The accuracy of the Immediate Test is superior to that reported for other short tests. On five populations of 100 or more adolescent and adult males, validated against standard individual tests, correlations of validity have run from $r = .83$ to $r = .91$. Test-retest reliability on an adult population of 137 was $r = .90$.

*The Immediate Test is available through the Sheridan Supply Company, P. O. Box 837, Beverly Hills, California. The writer wishes to acknowledge the following who have assisted in the construction of the Immediate Test: Robert Henkel, Richard Blum, Michael Bryant, Robert Strachan, Chauncey Martin, Victor Kirk and Herbert Sanderson.

Vocabulary as a Measure of Intelligence

Intelligence test constructors have concluded that the single most effective sub-test in individual test batteries is vocabulary. Terman and Merrill, and Wechsler, for example, in evaluating the adequacy of various items in their tests declare that from all points of evaluation that their vocabulary lists were the most satisfactory in all respects. These conclusions are the more remarkable since careful examinations of the Stanford Binet and the Bellevue vocabularies indicate that these lists have a number of very serious defects. Both lists of words were selected at random by the use of simple chance techniques, and on the Wechsler test there are only two words per mental year and only one word per mental year on the Terman list! Some words, such as *join*, and *regard* tend to be very difficult for the subject to define so that the examiner can decide whether the subject knows them. Some words are difficult to pronounce distinctly, i.e., *ochre*, *fur* and *roar*. Other words tend to be near-homonyms and the hapless subject may define *bacon* as *baking*, *roar* as *raw*, and *tint* as *tent*. This puts the examiner in the difficult position of saying, "I didn't say *tent*, I said *tint*."

Our central thesis is that since these two lists which appear to have many faults are nevertheless very valuable indicators of intelligence, that a new list, better selected and more adequately arranged, should form the basis of an acceptable short test of intelligence.

Construction of the Immediate Test

The plan for the construction of the Immediate Test was as follows: to obtain a pool of acceptable words; to find the difficulty level of these words; to select an adequate number of words at each level of intelligence; to evaluate the discriminative power of the words; and, finally, to evaluate the resulting list as a whole. Our first step was to ask the psychological staff of an Eastern reformatory to produce a list of words for the purpose of creating a new oral vocabulary list. These words, it was emphasized should:

1. Be neither too easy, nor too difficult.
2. Be easy to pronounce so that the subjects would "catch" them.
3. Be easy to define, preferably in one word.
4. Have but one common definition.

Some 300 words brought in by the collaborators were discussed and criticized in committee. Two hundred words were discarded and 100 were retained for trial. Three psychologists in the group administered the entire list of 100 words to five subjects each. For each word it was recorded whether the response was or was not correct, how many words were used to define the word, and whether the subject misunderstood the word. Following this, on the basis of discussion, fifty words were dropped, leaving fifty words as definite possibilities.

This procedure was repeated several times until a pool of 200 acceptable words in terms of our four-point criterion was obtained. They were then divided into two lists in approximate order of difficulty and each list was given to fifty subjects, young men in the age group 16-21, with known mental ages. We also gave each subject the Terman list. Correlations between the criterion test, the Terman list and our experimental lists were encouragingly high.

By re-analyzing the results of this administration, twenty words were eliminated and a list of eighty words was established. The new list was now given to 104 subjects for each of whom a Binet score was available and the 50% passing point established for each word. To our dismay we had a large concentration of words at the XI, XII, and XIII year levels, and another concentration at XVII and XVIII with very few words at XIV, XV and XVI. Another pool of acceptable words was established, this time attempting to concentrate on words at the XIV, XV and XVI

levels. About two years after the start we had eight words for each mental year from X to XX, but were not able to attain our original goal of twelve words per year level. By this time the constructor of this test transferred to an adult prison where a great many individuals had been given the Wechsler-Bellevue Test. The list was now tried on 100 subjects horizontally distributed on the I. Q. continuum. An item analysis and another difficulty analysis were made. Our item analysis showed that each word measured in the positive direction.

The method of item analysis was as follows: We established each word at a particular mental year level. This was the point of the 50% difficulty level of the word on the basis of the mental age, obtained by transposition from the Wechsler Efficiency Quotients. We were interested in the number of people with mental ages above the mental age level of the word, who failed it and those below the level of the word who passed it. The formula was Passed "low" plus failed "high" over the total number or $\frac{Pl + Fh}{H + L}$. Scores of .50 would indicate no discrimination. Scores below .50 indicated the word was valid. All of our words passed the criterion of below .50. At each level we dropped the two words that had the poorest discrimination, leaving the six best ones.

TABLE 1. APPROXIMATE DISTRIBUTION OF SCORES OF THE W-B, S-B AND IMMEDIATE VOCABULARY TESTS WITH REFERENCE TO MENTAL AGES.

MENTAL AGE	WECHSLER-BELLEVUE	STANFORD-BINET	IMMEDIATE TEST
20	36		66
19	31	25	60
18	29		54
17	27	23	48
16	25		42
15	22	19	36
14	20	16	30
13	17		24
12	15	14	18
11	13		12
10	10	11	6

The new list was given to a population of 300 subjects normally distributed as to intelligence who had Wechsler-Bellevue results. Correlations were run against the Verbal Weighted Score, the Performance Weighted Score, the Global Weighted Score, the Efficiency Quotient and the Intelligence Quotient. (Table 2.)

TABLE 2. PEARSON CORRELATIONS WITH IMMEDIATE TEST

TEST	*SUBJECTS*	*N*	*r*	*S.D.*
M.A. (OTIS)	Reformatory Inmates	50	.82	.05
" "	" "	"	.77	.06
TERMAN LIST	" "	"	.79	.06
" "	" "	"	.86	.03
M.A. (S.B. "L")	" "	104	.83	.03
WECHSLER-BELLEVUE I TWS	Prison Inmates	100	.89	.03
WEIGHTED SCORE-VERBAL	" "	300	.91	.01
" " Performance	" "	"	.86	.015
" " Total	" "	"	.90	.01
" " Efficiency Quotient	" "	"	.89	.015
" " Intelligence Quotient	" "	"	.90	.01
WECHSLER-BELLEVUE II WEIGHTED SCORE-TOTAL	" "	100	.84	.03
	Reliability			
TEST-RETEST	Prison Inmates	137	.90	.02

Administration

The Immediate Test is a vocabulary list of 66 words arranged in eleven groups of six words each. The groups represent Mental ages from X to XX.

The test is administered as a year scale, in the same manner as the Stanford Binet, with a basal and a terminal level. This allows testing to take place only within the individual's effective level of operating and gives maximum units of evaluation in a minimum time. The raw score can be converted into a mental age, or for adults directly into equivalent IQ's from a table imprinted on the test sheet. The test is administered in a brisk fashion. One scores plus or minus as one goes along. Long hesitations are marked minus, and the examiner gives the next word whenever the subject has given enough of the word to show he knows it.

Scoring

1. Scoring is soft, or lenient.
2. Each item is a member of some class, identification of the class is sufficient for credit.

Examples:

X	Zebra	*PLUS*	Animal. Striped animal. Like a horse.
		MINUS	I don't know. Bird.
XI	Nutmeg	*PLUS*	Spice. You put it in cakes. Food.
		MINUS	A tool.
XII	Bronchitis	*PLUS*	A disease. A sickness in the chest.
		MINUS	A horse.
XIII	Cathedral	*PLUS*	A church. A Catholic Cathedral.
		MINUS	A movie house.
XIV	Spigot	*PLUS*	A tap. A water faucet. A bung.
		MINUS	A spike. A tool.
XV	Kyack	*PLUS*	A boat. A canoe. Eskimo boat.
		MINUS	A house.
XVI	Fedora	*PLUS*	A hat.
		MINUS	I don't know.
XVII	Cranium	*PLUS*	The skull. The head. Part of the body.
		MINUS	A vegetable.
XVIII	Highboy	*PLUS*	A piece of furniture. Sort of a sideboard.
		MINUS	A tall man.
XIX	Metronome	*PLUS*	A time indicator. You practice the piano with it.
		MINUS	I don't know.
XX	Obsidian	*PLUS*	Rock. Volcanic glass.
		MINUS	I don't know.

In actual practice, the nature of the words is such as to produce either correct responses or inability to answer; rarely are minus responses given. Due to the fact that all words are nouns, and mostly of concrete objects, it is fairly easy to determine without a key which answers are correct. The Immediate Test has been administered by several examiners with the above directions without any difficulty.

Summary

In view of many deficiencies of current short scales, we have attempted the construction of an Immediate Test which is composed of a list of 66 carefully selected and screened words. This vocabulary test has acceptable statistical qualities, and in view of its convenience, shortness, and accuracy is presented as a workable clinical tool.

THE HILLSIDE SHORT FORM OF THE WECHSLER BELLEVUE

MILTON S. GURVITZ

Hillside Hospital

In the last five years, there have been many papers and literature dealing with Short Forms of the Wechsler-Bellevue test. None of these, however, exhaustively explored possible test combinations. The present study will attempt to discover the best possible two test Short Form of the Wechsler-Bellevue for relatively normal subjects and then test the applicability of this short form on substantial groups of patients of various nosologies to discover the extent to which they are applicable to pathological cases.

Method

Subjects for the experimental part of this study are believed to represent a substantially large enough group to make the results valid. The experimental group consists of 750 federal prisoners diagnosed as "without mental disease", 300 cases from Wechsler's standardizing population originally used for another study, 200 cases available to the author as the result of vocational guidance or experimental cases of individuals apparently "Normal", and 54 cases from Rapaport's Patrol. These total 1,304 cases.

To discover the best possible 2-test Short Form, each individual test in the experimental group had its sub-test introduced into a matrix. This matrix allowed every possible two-test combination for each individual Wechsler-Bellevue to be compared with the full scale rated score. The ten sub-tests allowed 45 possible combinations. When the 1,304 matrices were examined, and tabulated, they showed a wide distribution in the ability of two-test combinations to predict the final rated score. The lowest range, represented by the sub-test combinations of 8 and 7, predicted the final score perfectly in 101 cases out of a possible 1,304. The highest range was represented by the sub-test combination of 10 and 1 with 323 perfect predictions out of a possible 1,304. Other sub-test combinations which might be considered include: 10-2 with a score of 269, 7-1 with a score of 267, 7-4 with a score of 259, and 6-3 with a score of 227.

To further differentiate the best possible two-test short form, the five leading combinations of sub-tests were correlated with the full-scale rated score and their average deviation, irrespective of sign, computed from the same weighted score. The test combination which emerges as the best predictor will then be compared with the runner-ups on the three non-prison groups in the experimental population, with special emphasis on the two groups tested independently of the author, 300 cases from Wechsler's original population and 54 from Rapaport's control group. In addition, the effectiveness of the short form on various pathological groups will be evaluated by computing the average deviation from the full scale IQ on a large number of patients of varied diagnoses.

Results

Analysis of the results in Table 1 reveals that the test combinaion of 10-1 (information and digit symbol) retains its position as the best possible two test combination with both the highest possible correlation and the lowest average deviation.

Table 1. Correlation and A. D. of Certain Subtest Combinations With Full Scale Weighted Score

Subtests	*N*	*A.D.*	*r*
10–1	1304	5.9	.90
10–2	1304	7.0	.86
7–1	1304	7.2	.85
7–4	1304	7.4	.84
6–3	1304	7.9	.825

It is of course valuable to know whether the previous results were conditioned mainly by the large group of "normal" prisoners and whether the same results were reflected by non-prison groups and groups tested independently of the author. The 750 federal prisoners had an A. D. from full scale IQ of 6.0; the 300 cases from Wechsler's population an A. D. of 5.45; the 200 vocational guidance and experimental cases an A. D. of 6.2; and Rapaport's Patrol an A. D. of 5.9. The range appears to be rather small and it will be noted that some of the independent groups, notably the group from the original standardization population, showed even more promising results. The mean of the 1,304 experimental cases on the full scale IQ was 97.9. The comparable mean for the short form IQ was 98.7. The equivalence of the means is quite good and essentially the same for all practical purposes. The standard deviation for the full scale IQ was 14.6 and the standard deviation for the short form IQ 15.4. As might be expected from reducing the number of tests from 10 to 2, the standard deviation of the short form is somewhat larger.

Having established the sub-test combination of 10-1 (information and digit symbol) as worthy of consideration, it will be christened the "Hillside" Short Form after the hospital from which most of the clinical cases to be dealt with below were examined.

Thus far, we have dealt with relatively normal subjects. How will the Hillside Short Form stand up with pathological cases which can be expected to have a much greater intratest and intertest scatter? This question can be answered from Table 2. It would appear, that the Hillside Short Form does not appreciably lose its effectiveness when applied to psychopaths, alcoholics, and inmates of a Federal Prison Hospital.

TABLE 2. COMPARISON OF INTELLIGENCE QUOTIENTS OF SHORT FORM AND FULL SCALE WECHSLER-BELLEVUE ON VARIOUS PATHOLOGICAL GROUPS

Group	*N*	*Average Deviation Between Forms*
Total Schizophrenics	100	7.4
Hillside Schizophrenics	25	7.0
Rapaport's Schizophrenics*	75	7.5
Total Early and Pre-Schizophrenics	90	7.0
Hillside Early and Pre-Schizophrenics	45	6.80
Rapaport's Pre-Schizophrenics*	45	7.15
Total Depressions	62	7.7
Hillside Depressions	31	7.6
Rapaport's Depressions*	31	7.8
Total Hysterias & Neurasthenics	50	7.4
Hillside Hysterias and Neurasthenics	25	7.6
Rapaport's Hysterias and Nauresthenics*	25	7.2
Total Mixed Neurotics	38	7.35
Hillside Mixed Neurotics	19	7.5
Rapaport's Mixed Neurotics*	19	7.2
Total Obsessive Compulsives	34	6.75
Hillside Obsessive Compulsives	17	6.9
Rapaport's Obsessive Compulsives*	17	6.6
Criminal Psychopaths from a Federal Prison	200	6.1
Alcoholics, from the Bridgeport and Stamford Clinics of the Connecticut Commission on Alcoholism	50	6.3
A mixed group of mentally and physically ill prisoners from a Federal Penitentiary Hospital	250	6.1

*D. RAPAPORT, *Diagnostic Psychological Testing*, pp. 516-521.

When the Hillside Short Form is applied to Neurotics and Psychotics, there appears to be an appreciable loss of efficiency but when it is realized that the application of the Wechsler-Bellevue to such groups usually leads to an IQ underestimating the persons' potentialities plus the fact that in pathological groups the Hillside Short Form usually gives a higher IQ, we can postulate that the Hillside Short Form IQ will more often than not indicate more accurately the premorbid IQ than the Full Scale Wechsler-Bellevue IQ.

Another dimension yet to be explored is that of intelligence level. Does the Hillside Short Form perform equally well at all intelligence levels?

TABLE 3. COMPARISON OF FULL SCALE AND SHORT FORM INTELLIGENCE QUOTIENTS AT VARIOUS INTELLIGENCE LEVELS.

Group and Intelligence Level	*A.D.*
Experimental Group, Very Superior	4.75
Experimental Group, Superior	5.93
Experimental Group, Bright Normal	7.26
Experimental Group, Average	6.25
Experimental Group, Dull Normal	6.67
Experimental Group, Borderline	5.74
Experimental Group, Feebleminded	4.54

In general, it would appear that individuals at the ends of the scale have a better relation between full scale and short form than those around the middle. Greatest discrimination occurs at the level of Borderline and Feeblemindedness.

DISCUSSION AND CONCLUSIONS

The data seem to indicate that the Hillside Short Form of the Wechsler-Bellevue correlates quite adequately with the Full Scale IQ, much better than most of the other tests reported in the literature, and as well as the verbal and performance portions of the test as reported by Wechsler. There is a better equivalence between the short form and the full scale, than there is between the verbal and performance parts of the Wechsler-Bellevue since their *median* deviations are much larger than the *average* deviations reported in this study. Average deviations were used in this study because they adequately represented extremes. If we had used medians, the figures would have been much smaller, and the favorable difference between the results of this study and the median difference between verbal and performance as reported by Wechsler would have been enhanced.

It cannot be cautioned too strongly that the sole purpose of the Hillside Short Form of the Wechsler-Bellevue is to obtain an IQ. The diagnostic and clinical values of the test are totally abandoned in this shortened form. What then is gained? It usually takes about 70 minutes to administer the full scale Wechsler-Bellevue. The Hillside Short Form takes under ten minutes to administer. This allows the test to be administered when there would ordinarily be no time to administer a test at all. It is particularly adaptable to be included in the psychiatric mental status examination in place of the unstandardized questions so prevalent in the psychiatric manuals.

In any situation where a rapid estimate of intelligence must be made and adults or adolescents are dealt with, the Hillside short form will be of value. It should be emphasized that it will have a closer relation to the full scale IQ than a much longer *group test.*

Both tests are extremely objective in their scoring and no materials other than the standard test blank are needed. Administration is extremely simple:

1. Administer the information test, obtain weighted score.
2. Administer the digit symbol test, obtain weighted score.

3. Add weighted scores of 1 and 2.
4. Multiply the sum obtained in step 3 by *5*.
5. Use the product obtained in step 4 to obtain a Wechsler IQ using the same table as Full Scale IQ with due regard for the age of the subject.

Summary

Every possible two test combination of Wechsler-Bellevue sub-tests was examined to find which would make the best two test short form in 1304 non-pathological subjects. This proved to be Information and Digit Symbol with an r of .90 and an Average Deviation of 5.9 from the Full Scale IQ. This Hillside Short Form was tested with 874 pathological cases where it appeared that it was just as applicable to alcoholics, psychopaths and prisoners. On psychotics and neurotics the relationship to the actual Full Scale IQ was slightly reduced because the short form appeared to reflect to some degree the premorbid intellectual level. The Hillside Short Form is particularly effective at the lower intellectual levels. Administration takes but 10 minutes, and the short form should be especially valuable in estimating intelligence in psychiatric mental status examinations.

WECHSLER BELLEVUE SUBTEST SCATTER IN THE AFFECTIVE DISORDERS

SAMUEL WALDFOGEL AND WILLIAM GUY

McLean Hospital

Introduction

In recent years psychologists have been using subtest scatter on the Wechsler-Bellevue Scale increasingly as an aid to clinical diagnosis. However, the principles of scatter analysis for the most part are based on populations that have usually been small and unrepresentative in regard to such factors as age, sex, socio-economic status and education. Much additional work needs to be done on new populations and in various institutions. Only as these supplemental data are amassed can the psychodiagnostic value of scatter analysis on the Wechsler-Bellevue be properly evaluated.

The present report deals with the Wechsler-Bellevue performance of a group of patients with affective disorders. Little has appeared in the literature on this group of disturbances and the findings that have been reported are not in agreement. Rapaport et al.[6] report the results of four groups of depressive patients: psychotic, involutional, severe neurotic and neurotic who showed consistent impairment on the Performance subtests, the extent of impairment being roughly parallel to the severity of the depression.[1] Rabin[5], on the other hand, reports that manic-depressive patients do better on the Performance subtests than on the Verbal. In the same report Rabin described a Schizophrenic Index differentiating manic-depressives and schizophrenics. Gilliland, Wittman and Goldman[3] found that manic patients did slightly worse on most performance tests, but that in general their scatter was not significantly greater than the scatter of normal controls. However, only those patients who showed favorable test attitudes on the Elgin Test Reaction Scale were tested. This restricted their group to those that were least disturbed, which could easily account for the fact that they were not significantly different from normals in their test performance.

Methods

The patients in the present study were given Form I of the Wechsler-Bellevue Scale of Adult Intelligence as part of their diagnostic appraisal at the McLean Hospital. They were selected from three nosological groups: (1) manic-depressive depressed; (2) manic-depressive manic; and (3) reactive depression. The final diagnosis was the decision of the attending psychiatrist, usually after consultation with the combined staff at a case conference. Socio-economically the patients in all three groups were well above average in regard to wealth, class position and education. None of the patients in the present series had received any form of shock therapy prior to testing.

Results

The mean IQ's and mean weighted subtest scores for the three groups of patients are presented in Table 1 showing the number of patients in each group and their mean age. Table 1 reveals a good deal of similarity among the three groups. The mean IQ in all three is definitely above average despite the presence of severe psychopathology. This is probably a reflection of the patients' superior socio-economic level.

Only the manic group shows any appreciable difference between the scores on the Verbal Scale and the Performance Scale. However, if the subtests are consid-

[1]These patients also show considerable impairment on the Arithmetic and especially the Digit Span tests. However, this pattern is common to many other nosologic groups in Rapaport's series and does not differentiate the depressives.

ered in terms of their deviations from the Vocabulary scores, it is found that in general the Performance tests show much wider deviations than the Verbal test.[2] With the exception of Digit Span and Arithmetic the negative scatter is no greater than 2 on the Verbal subtests, whereas it exceeds 2 on all of the Performance subtests. While the manic patients show a greater degree of negative scatter than either of the depressive groups, the pattern for all three groups is similar. This pattern is similar to what Rapaport et al.[6] found with their depressive patients, but is the opposite of what Rabin[5] obtained with his manic-depressives.

In view of the differences in nomenclature and the general uncertainty of psychiatric classifications, it is difficult to make direct comparisons, but it seems that our manic-depressive depressed patients corresponded roughly to Rapaport's psychotic depressives and that our reactive depressives corresponded to his severe neurotic depressives. In comparing these groups it was found that our reactive depressives showed about the same degree of scatter as Rapaport's severe neurotic depressives. The sum of the negative Vocabulary Scatter of the five Performance subtests averaged 15.9 for our patients and 13.4 for his. However, our manic-depressive depressed patients had a mean total scatter of only 16.6 on the same tests compared to a mean total of 25.6 for Rapaport's psychotic depressives. The latter figure was similar to the mean total of 22.6 which we found for our manic-depressed manic patients. In other words our manic-depressive manic patients resembled Rapaport's psychotic depressives more closely than did our manic-depressive depressives.

The severity of the depression at the time of testing could easily have influenced the amount of negative scatter and there is no way of knowing to what extent our depressive group was comparable to Rapaport's in this respect. The difference between them might be accounted for on this basis. With the manic patients, however, the overt symptomatic picture is in striking contrast to the depressive patients, and their similarity in performance to the depressives requires some explanation.[3] The greatest impairment, it will be recalled, was on the Digit Span, Arithmetic and the five Performance tests. The two former tests place an especially high premium on attention and concentration, the latter group in addition places great emphasis on speed of performance. It is not difficult to understand how either the retardation of the depressive or the excitement and distractibility of the manic could have a detrimental effect on the performance of these tests. Though the net result is the same, the underlying determinants could thus be quite different. In any event, it is evident that this subtest pattern is not unique to depressive patients.

One of the chief criticisms leveled against scatter analysis as it is being used here is that the weighted subtest scores, unlike the IQ, are not corrected for age[1, 2, 4]. Since it is known that some subtests are affected more by age than others, it can readily be seen that the subtest pattern of any nosologic group under consideration may be influenced by the age of the members of the group. To determine if age played any part in influencing the subtest patterns of our patients, we divided them into two equal groups on this basis. Since the subtest patterns of our three nosological groups paralleled one another we pooled them for this comparison. The scores of the patients in the older half of the total group were then combined and compared with the scores of the patients in the younger half. The mean age of the two resulting groups and their mean scores are given in Table 1.

---

[2]We are following the practice of Rapaport et al. (6) in using Vocabulary Scatter. This method of scatter analysis seems especially well suited to clinical diagnosis. The weighted Vocabulary score is treated as a baseline. Any subtest weighted score that is higher than the Vocabulary is regarded as positive scatter; any subtest weighted score below Vocabulary is regarded as negative scatter.

[3]The implication is that on the basis of scatter alone one would be hard put to distinguish between manics and depressives. This is a dramatic reminder that scatter analysis can never be done in a purely mechanical fashion; and that qualitative as well as quantitative features of performance must be taken into account.

TABLE 1. MEAN I.Q.'S AND WEIGHTED SUBTEST SCORES ON THE WECHSLER-BELLEVUE SCALE FOR PATIENTS WITH AFFECTIVE DISORDERS, AND ALSO BY AGE GROUPS

	No. of cases	Mean Age	Full Scale	Verb. Scale	Perf. Scale	Vocab.	Inf.	Comp.	Digits	Arith.	Simil.	P. Arr.	P. Comp.	B. Des.	O. Ass.	D. Sym.
Manic-depressive manic	16	35.3	105.8	109.2	101.1	13.5	12.6	11.3	8.4	9.7	11.1	8.5	9.7	9.4	8.4	8.9
Manic-depressive depressed	21	46.0	117.3	117.2	115.2	13.5	12.5	12.1	9.9	11.9	12.6	9.3	11.0	11.1	9.8	9.5
Reactive depressiln	21	39.5	118.1	117.2	116.5	14.1	13.7	12.8	10.2	10.9	12.2	10.3	11.5	11.4	10.7	10.7
Younger Patients	29	31 1	114.6	114.9	112.0	13.4	12.6	12.0	9.7	11.7	12.5	10.4	11.5	11.7	10.0	11.0
Older Patients	29	50.0	114.1	114.6	112.1	14.0	13.3	12.1	9.4	10.0	11.6	8.3	10.3	9.7	9.5	8.5

While the IQ's for the two groups are virtually identical, the subtest results clearly reveal the relative superiority of the younger group on the Performance tests. The negative scatter on the Performance tests is almost twice as great for the older group as for the younger. The mean total scatter for the Performance tests is 23.8 and 12.3 respectively for the two groups. The difference between these totals is highly significant from a statistical standpoint: $P = < .001$. On the basis of these results it appears that age adds a considerable amount to the extent of negative scatter in our patients. If the age factor were held constant, we would expect the subtest pattern to be somewhat less sharply defined and hence to have less diagnostic significance than seemed to be the case initially.

One further analysis of the data was made. This was in connection with Rabin's Schizophrenic Index[5] for differentiating schizophrenics from manic-depressives.[4] The index was calculated for our three groups and the results are presented in Table 2 where they are compared to Rabin's findings.

TABLE 2. THE SCHIZOPHRENIC INDEX: A COMPARISON OF MCLEAN HOSPITAL PATIENTS WITH THOSE STUDIED BY RABIN.

Group	Mean Schizophrenic Index
Rabin's schizophrenics	1.27
Rabin's "manic-depressives"	1.08
McLean manic-depressive manic	1.21
McLean manic-depressive depressed	1.16
McLean reactive depression	1.17

It can be seen that values for the McLean patients fall midway between the values for Rabin's schizophrenics and his manic-depressives. Further it was found that 31% of the McLean patients fell above Rabin's mean schizophrenic score of 1.27. On the basis of these findings it appears that the Schizophrenic Index has little value for differentiating the affective disorders from schizophrenia.

DISCUSSION

The results of the present study confirm the previous findings of Rapaport[6] that depressive patients reveal a characteristic subtest pattern on the Wechsler-Bellevue Scale. However, it was found that his pattern was less specific than they had assumed since it also appears in manic patients. It was further shown that this pattern was more marked among the older patients in our groups and seemed to be a function of age as well as of psychopathology. The influence of age on the subtest pattern serves to reduce its diagnostic sensitivity. On the basis of these findings it can only be concluded that the diagnostic value of the Wechsler-Bellevue Scale is limited in the area of the affective disorders. While it may serve as an aid to diagnosis in this area it can hardly be regarded as a reliable indicator, either for the purpose of differentiating the affective disturbances from other forms of psychopathology, or for distinguishing between manics and depressives.

---

[4] $\text{Schizophrenic Index} = \dfrac{\text{Information} + \text{Comprehension} + \text{Block Design}}{\text{Digit Symbol} + \text{Object Assembly} + \text{Similarities}}$

Conclusions

1. Patients with affective disorders show similar subtest patterns on the Wechsler-Bellevue Scale regardless of whether they are classified as manic or depressive.

2. This pattern is characterized primarily by large negative Vocabulary Scatter on the Performance Scale.

3. The above pattern, however, proved to be in part a function of age since it was more marked among the older patients in our groups. Since the pattern fluctuates with age its diagnostic reliability is thereby diminished.

4. Rabin's Schizophrenic Index which allegedly differentiated manic-depressives from schizophrenics did not prove to be a reliable indicator with our patients.

5. At best Wechsler-Bellevue scatter can only be used as a rough aid to the diagnosis of affective disorders. By itself it cannot be relied upon to yield accurate diagnoses.

Bibliography

1. Barnett, I. The use of Z scores in equating the Wechsler-Bellevue subtests. *J. clin. Psychol.*, 1950, 6, 184-188.

2. Foster, A. Age and the Wechsler-Bellevue scattergraph. *J. clin. Psychol.*, 1947, 3, 396-397.

3. Gilliland, A. R., Wittman, P., and Goldman, M. Patterns and scatter of mental abilities in various psychoses. *J. gen. Psychol.*, 1943, 29, 251-259.

4. Jastak, J. Problems of psychometric scatter analysis. *Psychol. Bull.*, 1949, 46, 177-197.

5. Rabin, A. I. Differentiating psychometric patterns in schizophrenia and manic-depressive psychosis. *J. abnorm. soc. Psychol.*, 1942, 37, 270-272.

6. Rapaport, D., Gill, M., and Schafer, B. S. *Diagnostic psychological testing*, Vol. I. Chicago: The Year Book Publishers, 1945.

A TEST-RETEST EVALUATION OF THE WECHSLER FORMS I AND II WITH MENTAL DEFECTIVES*

WILLIAM HAYS AND BERNARD SCHNEIDER

Lincoln State School and Colony

Introduction and Statement of Problem

The Wechsler-Bellevue Intelligence Scale is the most used single measure of intelligence in the adult and adolescent. In particular, it has found wide use in the evaluation of borderline and high grade defective adults, for which purpose its features of interest, brevity, and "global" approach to intellectual ability are important.

However, the problem of test and retest with such an instrument becomes puzzling in the clinical situation. This is true particularly in the precommitment and post-commitment examination of suspected mental defectives in which a more or less arbitrary legal IQ limit has been set to classification as mentally defective. Similarly, it is oft-times desirable in the institutional setting to seek comparable intellectual evaluations within a relatively short space of time. In such cases, we should expect some variation between test scores due to the practice effect or general "carry over" from one testing to another, as well as to motivational and chance factors.

Form II of the Wechsler Scale (3) was published in 1946 with the stated intention of meeting this need. However, little material has been furnished to date to indicate the equivalence of Form II to Form I, and the amount of variation we should be led to expect in various populations between the initial testing and the retest with the same and with different forms of the test. Gibby (2) has reported that the two forms of the Wechsler Scale agree very well in regard to psychoneurotic subjects. However, he found many disagreements in subtest performance between the two forms. Brown (1) found that the two forms of the Wechsler are not fully comparable in regard to distortion of the picture of intertest variability.

While the important problem of interest variability between the two forms of the scale is, unfortunately, beyond the range of this report, the authors are concerned with the immediate and practical problem of the evaluation of the effect of test upon retest with the two forms of the Wechsler. Obviously, we should expect some practice effect from the readministration of the same form of the scale; likewise, we might expect some general "carry over" for the different, yet materially similar Forms I and II. We should not expect the order of presentation of different forms of the test to affect the amount of increase obtained, however. In such an evaluation of the "carry over" from one testing to another, we should also expect time interval between tests to be a factor affecting score increase as well.

Thus, the problem of this study is twofold: (1) a comparison of the obtained test-retest differences between the four possible different orders of administration of the Wechsler I and II; and (2) a comparison of test-retest differences for similar groups after the arbitrarily chosen time intervals of 2, 4, 6, and 8 weeks.

Procedure

Eighty high grade and borderline defective patients of Lincoln State School and Colony were utilized for this study. These patients were drawn from the familial and undifferentiated classifications exclusively. No patient was utilized who presented any observable neurological involvement or sensory incapacity. Chronological age ranged from 15 to 28 years. No subject was included for whom there was record of any previous Wechsler Scale administration. IQ's, based upon previous Binet Tests, ranged from 50 to 78.

A random assignment of subjects was made to one of 16 groups, each containing five individuals. These groups then were given the Wechsler Scale in one of the four

*From the Psychological Laboratory of Lincoln State School and Colony.

test-retest sequences. In addition, each sub-group within the four larger groups was allowed a testing interval of either 2, 4, 6, or 8 weeks, respectively.

Table 1 presents in graphic form the design of this experiment. From this design it can be seen that each of the 16 sub-groups of five individuals fulfilled one condition of test presentation and one condition of time interval, thus making possible an analysis of variance for both order of presentation and time interval, as well as for the interaction of one upon the other.

TABLE 1. MEAN INCREASE FOR EACH CONDITION OF TIME AND ORDER

Test Order	*Time Interval in Weeks* 2	4	6	8	Mean
I–I	9.6	7.1	5.4	8.1	7.57
I–II	−3.4	6.8	−1.1	−.6	0.42
II–II	−1.2	7.2	7.6	3.8	4.35
II–I	5.6	5.7	3.6	3.1	4.51
Mean	2.66	6.71	3.87	3.60	

Two experienced test administrators were employed, and administration error was minimized through allowing both test and retest on one subject to one administrator. An equal number of individuals was examined in each time and order category by each examiner. A rigid standard scoring procedure was employed, and scoring was cross-checked by each examiner.

In order to avoid artificial differences due to chronological age and item placement on different test forms, all differences were measured in terms of weighted scores. A full administration of eleven subtests for both forms was employed, with proration for the Verbal Score to include Vocabulary, in order to obtain full scale weighted scores. In accordance with the suggestion of the test manual, the correction was made for verbal, performance, and total weighted scores on the Wechsler II before differences were computed.

RESULTS

An initial assumption underlying any analysis of variation in test scores is that of homogeneity of the respective groups with regard to the initial level of ability before the comparison of retest differences. Therefore, the total initial weighted score of the individuals was utilized as a comparative measure of initial ability for the various subgroups. These groups were then subjected to the Bartlett X^2 test for homogeneity, with the result that the sixteen random subgroupings were found to be homogenous in respect to initial test ability, since the X^2 was significant at only the 30% level of confidence.

Results in terms of weighted score differences were treated then according to analysis of variance from a 4 x 4 table with multiple classification. A null hypothesis of no difference in test-retest increase for both order of presentation and for time interval was proposed. Under these conditions an F test showed order of presentation to be significant at less than the 5% level of confidence. However, F for time interval and for interaction both proved insignificant at greater than the 5% level.

The mean increase in weighted score for order I-I was 7.57 points; for order I-II, 0.42 points; for order II-II 4.35 points; and for order II-I 4.51 points.

Upon the null hypotheses that there should be no difference between test I presented first and test II first, no difference between retest with I and retest with II, and between retest with the same and with different tests, an analysis of variance was set up with the four grouped orders of presentation, utilizing a 2 x 2 factorial design of the type ABBA. An F test here showed the difference between increases when test I was utilized first and test II first to be highly insignificant. However, it was found that an F for difference in increase between retest with I and with II was highly significant at much less than the 1% level of confidence. Similarly, an F for the like retest versus the unlike retest was much lower than the 1% level, as expected.

While we should expect a considerably greater increase for test order I-I than for order I-II, we should also expect a greater increase for order II-II than for order II-I. This, however, is not the case. The *t* test was significant at less than the 1% level of confidence for differences in increase between any two orders of administration, with the exception of II-II versus II-I, which was highly insignificant.

Mean verbal increases were each found to be less than two weighted points. Mean performance increases were found to be 5.5 points for order I-I, −.05 points for I-II, 3.93 points for order II-II, and 2.2 points for order II-I.

A similar 4 x 4 replicated analysis was made separately for total verbal and for total performance items, once again utilizing a null hypothesis of no significant difference for either time or order of presentation of the two forms. F tests were insignificant for order, time, or interaction for the verbal totals. Nevertheless, on performance totals, F was insignificant for time and interaction, but significant at less than the 5% level of confidence for order of presentation.

The failure of the like retest versus the unlike retest to function according to expectation led the authors to formulate the hypothesis that the correction made for the Wechsler II scale might be a source of the variation from expectancy in these cases.

Employing uncorrected Wechsler II scores, the mean increase for order I-I was found to be 7.57 points; for order I-II, 2.42 points; for order II-II 4.35 points; and for order II-I 2.51 points.

Therefore, on the null hypothesis of no difference in increase in time interval or order of presentation, a 4 x 4 replicated analysis of variance was set up utilizing the data with uncorrected Wechsler II scores. Here, the F test showed time, order of presentation, and interaction each to be insignificant at greater than the 5% level of confidence.

Discussion

The results of these analyses would seem to indicate that there is insufficient change in weighted scores over a period of time ranging from two to eight weeks to be considered significantly different from chance variation. However, the order of presentation does seem to effect the amount of weighted score increase considerably. We are led to expect that similar tests will yield higher increases on retest than dissimilar tests, but, while this is true in the case of test I followed by test I as opposed to test I followed by test II, we find that almost equal increases occur between test II followed by test II, and test II followed by Test I.

Similarly, if the two forms of this test are truly equivalent, we should expect similar and insignificant increases between the readministration of Form I and of Form II. Contrary to expectation, a significant difference exists between the increases obtained from the readministration of the two forms. Once again, if the two forms be equivalent there is no reason to expect a significant difference between one order of presentation of retest with a different form and another order. Such difference was found to exist, however.

It will be noted that with high grade and borderline defectives the correction term offered for Form II has a differential effect upon retest increase, dependent upon order of administration of the different forms. For example, if raw increase for the order of presentation I-II is four weighted points, correction of Form II reduces this to two points. However, if increase for order II-I is four points, correction will make this difference six points.

When it is considered that the correction offered by Wechsler for the weighted scores of Form II has the effect of depressing increase upon retest with Form II following Form I, and of enlarging increase upon retest with Form I following Form II, a hypothesis might be extended that the correction factor for Form II is accountable for some of the variation in increase found. Furthermore, the drop in significance of differences between orders of administration as indicated by the analysis of uncorrected scores might serve to support this contention. This might also be borne out

by the consideration that the correction penalizes the performance items most heavily, these items being the source of considerable variation in increase for our groups It is interesting to note that the mean difference for uncorrected full weighted scores for test order I-II and for II-I agree closely both with each other and with the mean difference of two points reported by Wechsler (3) from preliminary comparisons.

While there is no indication made of the population employed in the preliminary comparisons of Forms I and II by Wechsler, it may be assumed that they were drawn largely from the average intelligence group, for which a large portion of the correction term is either not employed or else is a positive rather than a negative correction.

We are still at a loss to explain the significant difference in increase between readministration of I and of II. However, any explanation which could be made from our data would have little meaning without some estimate of the relative difficulty of the two forms in regard to some outside criterion, and the consequent amount of "carry over" which we should be led to expect from the readministration of the same test to defective subjects.

Barring faulty randomization of subjects and the necessarily uncontrolled factor of relative motivation in test performance, these results lead the authors to cast some doubts as to the present equivalence of Form II of the Wechsler Scale to Form I, in the matter of test-retest over relatively short intervals with high grade defective subjects. They are led to question in this respect the validity of the correction term given for Form II as applied to defective subjects, particularly with respect to performance and total items.

Summary and Conclusions

1. Eighty high grade defective and borderline patients were assigned to groups of five and were given the Wechsler Forms I and II, each group being given one of the four possible orders of test-retest, and being allowed either 2, 4, 6 or 8 weeks between tests.

2. Analysis of variance revealed no significant difference in weighted score increase for time and interaction. Order of presentation showed significant differences.

3. While there was no significant difference between increase when Form I was presented first and Form II first, significant differences were found between retest with Form I and Form II. Like retest versus unlike retest also showed highly significant differences.

4. Differences in mean increase were significant for any two orders of administration, with the exception of II-II versus II-I.

5. Verbal totals showed no significant score increase for order, time, or interaction.

6. Performance totals showed no significant score increase for time or interaction, but were significant for order of presentation.

7. An analysis of total score increases with uncorrected Wechsler II totals showed no significant differences for time, order of presentation, or interaction.

8. The conclusion was reached that, as they now stand, Forms I and II of the Wechsler Scale are not fully equivalent in the matter of test-retest increase with mental defectives. However, it was felt that the correction factor may be inappropriate for defective subjects and thus accountable for some of the variation observed.

References

1. Brown, M. N. A critique of the Wechsler-Bellevue system of Weighted Scores. *J. clinical Psychol.*, 1949, 5, 170-173.
2. Gibby, Robert G. A preliminary survey of certain aspects of Form II of the Wechsler-Bellevue Scale as compared to Form I. *J. clin. Psychol.*, 1949, 5, 165-169.
3. Wechsler, D. The Wechsler-Bellevue Intelligence Scale. Form II Manual for administering and Scoring the Test. New York: *Psychological Corporation*, 1946, pp. v-vi.

A STUDY OF THE DIFFERENTIAL RESPONSES ON THE VOCABULARY SUB-TEST OF THE WECHSLER-BELLEVUE INTELLIGENCE SCALE*

CHALMERS L. STACEY AND BERNARD PORTNOY

Syracuse University

Introduction

The authors, using a suggestion in an article by Reva A. Gerstein[1] as a point of departure, conducted a study intended to discover some relationship between the use of language and the underlying cognitive process[4]. That is, instead of merely determining whether a definition of a word given on a vocabulary test is correct or incorrect, the study aimed at differentiating the level of definition credited as correct, and to see whether groups at different intelligence levels would perform differentially.

Reichard, Schneider, and Rappaport[3], using the Goldstein-Weigle[2] Color Sorting Test as an instrument to study concept formation in children, concluded that three methods of forming concepts exist: (a) the concretistic, (b) the functional, and (c) the conceptual. Gertstein assumed the same three methods used in solving a performance problem may be used in classifying definitional levels produced by verbal stimuli, and suggested the following classification: (a) the concretistic or descriptive method, which is considered the most primitive level of verbalized definition, e.g. "An apple is red.", (b) the functional or usage method which is a more complex method of verbalization, e.g. "A bicycle is something you ride.", and (c) the categorical or conceptual method which represents an abstract method of definition. This type of definition is considered more complex than the preceeding two, e.g. "A donkey is an animal.".

In the author's previous study, the responses of two groups of children, one classified as mentally defective and the other as borderline, were analyzed in terms of total I.Q. on the Wechsler Intelligence Scale For Children, and as to the proportion of their vocabulary responses that fell into the descriptive, functional, or categorical categories. Contrary to what was expected, a significantly greater proportion of descriptive responses were found to occur in the borderline group than in the mental defective, and a significantly greater proportion of functional responses were found to occur in the mental defective group than in the borderline group. No significant difference was found among the categorical responses, though the borderline group did have greater absolute proportion. These results left in doubt the fact that the functional response was of a higher level than the descriptive. The purpose of the present study is to determine whether these results would be obtained in an adult sample.

Method

Three groups of adults were used as subjects. They were classified as moron, borderline and dull-normal according to the Wechsler-Bellevue Intelligence Scale. Their characteristics can be seen in Table 1. The responses of each subject to the words in the vocabulary section of the Wechsler-Bellevue Intelligence scale were recorded verbatim. Each response was entered on a chart under the appropriate category, the descriptive, the functional, and the categorical. Only nouns were used. The total number of words in each category was computed. The proportion of each type of response (descriptive, functional, and categorical) to the total number of responses was then calculated for each of the groups. Critical ratios were computed between nine pairs of proportions. The results are presented in Table 2.

*This study was done at the Syracuse State School, Syracuse, New York. The writers wish to thank Dr. B. S. Bisgrove, Senior Director, for his cooperation.

TABLE 1. PERFORMANCE OF MORON, BORDERLINE, AND DULL NORMAL SUBJECTS ON THE WECHSLER-INTELLIGENCE SCALE

	N	Chronological Age (in years)		Verbal Quotient		Performance Quotient		Total I. Q.	
		Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. S.
MORON (I.Q. 50-65)	35	18.18	2.33	60.77	3.52	64.57	6.53	58.37	4.01
BORDERLINE (I.Q. 66-79)	33	18.46	2.90	69.00	4.95	82.15	5.48	72.64	2.93
DULL-NORMAL (I.Q. 80-90)	24	19.68	2.86	77.71	6.34	92.17	7.34	83.12	2.92

TABLE 2. FREQUENCY AND PROPORTIONS OF LEVELS OF RESPONSE TO THE VOCABULARY SECTION OF THE WECHSLER-BELLEVUE INTELLIGENCE SCALE.

	Dull Normal			Moron			Borderline			Dull Normal	
Category	Frequency	Proportion	C.R.*	Frequency	Proportion	C.R.*	Frequency	Proportion	C.R.*	Frequency	Proportions
Descriptive	50	.24	.75	53	.21	.56	64	.23	.26	50	.24
Functional	86	.40	.93	119	.47	.93	120	.43	.67	86	.40
Categorical	76	.36	.89	80	.32	.49	93	.34	.45	76	.36
TOTAL	212			252			277			212	

Results

1. There was no significant difference between the proportions of descriptive responses among the three groups. That is, between moron and dull-normal, moron and borderline, and dull-normal and borderline. However, there was a trend for the proportions of descriptive responses to increase as we go from moron through the borderline to the dull-normal levels of intelligence.
2. There was no significant difference between the proportions of functional responses among the three groups. However, there was a trend for the proportions of functional responses to decrease as we go from moron through the borderline to the dull-normal levels of intelligence.
3. There was no significant difference between the proportions of categorical responses among the three groups. Again there was a trend for the proportions of categorical responses to increase as we go from moron through the borderline to the dull-normal levels of intelligence.

A comparison was then made between the results of our adult moron sample and child mental defective or moron sample, and the adult borderline sample and the child borderline sample. The following are the results of this comparison, as may be seen in Table 3.

4. In the two younger age groups, the highest proportion of responses fell in the functional category, with the descriptive coming next, and the categorical last. In the older age group, the proportion of functional responses was still highest, but the categorical came second and the descriptive last.
5. There was a significant difference in the proportions of descriptive responses between the adult moron and child moron samples; and between the adult borderline and child borderlne samples. In both cases, the adult groups had the greater proportions. In all cases, the differences were significant at the one percent level.
6. There was a significant difference in the proportions of functional responses between the adult moron and child moron samples; and between the adult borderline and child borderline samples. In both cases, the younger groups had the greater proportions. In all cases the differences were significant at the one percent level.
7. There was a significant difference in the proportions of categorical responses between the adult moron and child moron samples; and between the adult borderline and child borderline samples. In both cases, the older group had the greater proportions. In all cases the differences were significant at the one percent level.

Summary

The responses of three groups of adults, one moron, one borderline, and one dull-normal were analyzed in terms of total IQ on the Wechsler-Bellevue Intelligence Scale and as to the proportion of their vocabulary responses that fell into the descriptive, functional, or categorical categories. In contrast to what was obtained with two groups of children[4], no significant differences were found among the proportions of the three types of responses among the three adult groups, though the same trend was seen to emerge. When the two age groups were compared, a significantly greater proportion of descriptive and categorical responses were found to occur in the older age group than in the younger, and a significantly greater proportion of functional responses in the younger age group than in the older age group. The Reichard, Schneider, and Rappaport[3] study using performance tests stated that children think at a descriptive or concrete level up to about the age of eight. Between the ages of eight and nine the functional level is reached, and at eleven years of age the conceptual level is reached. Thus, we would expect the proportion of

TABLE 3.—FREQUENCY AND PROPORTION OF LEVELS OF RESPONSE FOR ADULT AND CHILDREN'S GROUPS TO THE VOCABULARY SECTION OF THE WECHSLER - BELLEVUE INTELLIGENCE SCALE.

Category	Child - Moron		Adult - Moron			Child - Borderline		Adult - Borderline		
	Frequency	Proportion	Frequency	Proportion	C.R.	Frequency	Proportion	Frequency	Proportion	C.R.
Descriptive	10	.054	53	.21	5.00	38	.144	64	.23	2.73
Functional	166	.897	119	.47	11.03	212	.803	120	.43	9.74
Categorical	9	.049	80	.32	7.71	14	.053	93	.34	9.35
TOTAL	185		252			264		277		

functional and categorical responses to be higher in the older age group, and the proportion of descriptive responses higher in the younger age group. These conclusions, in our study, were substantiated only in reference to the conceptual level. The difference might be due to our sample of below normal intelligence, or to the fact that Reichard, Schneider, and Rappaport's study was of a performance type, and ours was not. Further research in the field using subjects of normal intelligence would be quite illuminating regarding this question.

References

1. Gerstein, Reva A. A suggested method for analyzing and extending the use of Bellevue-Wechsler Vocabulary responses. *J. consult. Psychol.*, 1949, 13, 366-370.
2. Goldstein, K. and Scheerer, M. Abstract and concrete behavior; an experimental study with special tests. *Psychol. Monog.*, 1941, No. 2.
3. Reichard, S., Schneider, M., and Rappaport, D. The development of concept formation in children. *Amer. J. Orthopsychiat.*, 1944, pp. 156-161.
4. Stacey, Chalmers L., and Portnoy, B. A study of the differential responses on the vocabulary sub-test of the Wechsler Intelligence Scale for Children. *J. clin. Psychol.*, 1950, 6, 401-403.

THE EFFECT OF LATERALITY LOCALIZATION OF FOCAL BRAIN LESIONS ON THE WECHSLER-BELLEVUE SUBTESTS[1]

A. LLOYD ANDERSEN

Nevada State Department of Health

INTRODUCTION

In an earlier article by Andersen (1) evidence was presented that the Wechsler-Bellevue Deterioration Quotient and the Reynell Index were differentially effective in detecting surgically authenticated focal lesions in the dominant and non-dominant, or recessive, cerebral hemispheres. This finding naturally suggests that dominant and non-dominant hemisphere brain damage affects the Wechsler-Bellevue subtests differentially. The present paper deals with this problem.

The cases used in the present study were fifteen cases of focal brain damage to the dominant cerebral hemisphere and fifteen cases of focal brain damage to the non-dominant hemisphere. All subjects were male patients at the Minneapolis Veterans Administration Hospital. These thirty cases were the first fifteen suitable cases of each laterality group collected over a two-year period. The total group of fifty-five cases is described in the earlier paper. The twenty-five cases not included in the present analysis were reserved as a check or cross-validation group for an index of laterality localization.

PROCEDURE

It was felt that the best approach to the study of the dominant hemisphere (hereafter DH) and non-dominant, or recessive, (NDH) groups would be to study the total pattern or profile differences obtaining between the two groups. For this purpose it was desirable to render each subtest weighted score within each case in terms of some base which reflected the general level of all weighted scores constituting that particular case. Thus, each weighted score is expressed in terms of a base characteristic of the particular individual. In this way scores are rendered relative to the base (a characteristic of the individual) rather than relative to the standing of the standardization group for that subtest. This procedure renders the subtest scores more directly comparable from case to case and, depending upon the base selected, may partially rule out the effect of extraneous variables upon the profiles.

Several bases have been proposed by investigators studying profile differences between various clinical groups on Form 1 of the Wechsler-Bellevue. Among these have been the total weighted score employed by Wechsler (4, pp. 148-153) and the Vocabulary score which Rapaport (3) used extensively. However, a preliminary survey of the data seemed to indicate that a more satisfactory base for this particular problem would be the sum of the Information and Block Design subtests. Jastak (2) has pointed out that "the use of any reference point for the measurement of a trait is impractical if the trait in question has a variable effect on that reference point". In this case, the base (reference point) had to be dependable as possible in either laterality group, being influenced by brain damage in a consistent, rather than variable, manner. A preliminary rank analysis showed that both of the subtests (Information and Block Design) had relatively low rank variability in each laterality group and that variability was closer to equal between the laterality groups for these subtests than for any other pair of subtests. The Digit Span and Digit Symbol subtests were excluded from the study at this point because they appeared to be extremely sensitive to both DH and NDH brain damage and, in addition, were extremely

[1]Abstracted from a Ph.D. thesis presented to the University of Minnesota, June, 1949. This paper is published with the permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed nor the conclusions drawn by the author.

variable in both groups. The Total Verbal and Total Performance weighted scores were included in the analysis since they appeared differentially sensitive and dependable.

Having decided upon Information plus Block Design (weighted scores) as the most satisfactory reference point, all subtest weighted scores were converted to scores depending upon their position relative to the base value by the formula

$$(1) \quad \text{V score} = (\text{Info} + \text{BD}) - 2X + 10$$

where X is the subtest score being converted. Since it was desired that the scores vary around the base as a mean, it was necessary either to divide the base by two or multiply the subtest score by two in formula (1). The latter procedure seemed simpler. The addition of ten eliminated the majority of negative V scores throughout the analysis.

Use was made of the Total Verbal and Performance weighted scores by converting them to V scores by the formula

$$(2) \quad \text{Vv6} = \frac{3\,(\text{Info} + \text{BD}) - \text{v6}}{3} + 10$$

$$(3) \quad \text{Vp5} = \frac{5\,(\text{Info} + \text{BD}) - 2\text{p5}}{5} + 10$$

where v6 is the Total Verbal weighted score, p5 the Total Performance, Info the Information, and BD the Block Design weighted scores.

Because the Total Verbal V score should be fairly comparable to the subtest V scores, several adjustments were made. In formula (2) the V base (Info + BD) is multiplied by three to adjust it to a level comparable to the Total Verbal score based on six subtests. The difference between these values will then be three times as large, on the average, as the subtest V score based on sums the equivalent of only two subtests. Hence, the divisor of three. Comparable adjustments may be observed in formula (3).

The distributions of V scores for the two groups were then studied by the following methods:

1. The significance of the differences between the means of the DH and NDH groups for each of the eleven variables was checked with the t-test.
2. The values, within each case, of all possible sums of pairs of V score variables and of all possible differences between pairs were calculated.

Thus, the values of the sums of V Info and all other V score variables, taken one at a time, and of V Comp and all other variables, and so forth, were obtained for each case. The same procedure was followed in obtaining the differences between pairs. This resulted in fifty-five two-variable values of sums and fifty-five of differences. The significance of the differences between the means of these two-variable values in the DH and NDH groups was checked with the t-test.

Results and Discussion

The means and test of significance for the eleven single V scores (First Order) of subtests and Totals are given in Table 1. Tables 2 and 3 show the more significant sums of pairs and differences of pairs of V scores, respectively. No pairs are entered in Tables 2 and 3 whose means in the DH and NDH groups did not differ at the five percent level of confidence.

Table 1 indicates that four of the five verbal subtests and the Total Verbal variable are significantly higher for the DH group than for the NDH group. A study of formulae (1), (2), and (3) will indicate that this means that the DH brain damage group shows greater losses of verbal abilities than does the NDH brain damage group. On the other hand, the differences between performance subtests are all in the opposite direction, indicating relative greater loss in these functions among the NDH group. Study of Tables 2 and 3 will show that these general differences remain consistent in the significant two-variable combinations. All of the two-variable sums shown in Table 2 show differences in favor of the dominant hemisphere group, with only one exception; all of the variables represented are from the verbal portion of the Wechsler. The two-variable differences shown in Table 3 are just as consistent in this regard; all combinations with verbal variables to the left of the minus sign show differences in favor of the dominant hemisphere, while the last four two-variable differences (with performance variables to the left of the minus sign) show differences in favor of the non-dominant hemisphere brain damage group.

The differences seen in these tables suggest that the patient with dominant hemisphere brain damage forgets *what* to do, while the patient with non-dominant hemisphere damage forgets *how* to do. Observations and inquiries during testing seem to bear this out. While both groups may fail a test task, they apparently do so for different reasons. Thus, a dominant hemisphere patient may fail an arithmetic problem because he has forgotten portions of the multiplication tables, while the non-dominant hemisphere patient fails because he cannot decide what basic processes must be employed in solving the "story problem", though he remembers his "sums" and multiplication tables quite well. A fairly satisfactory analogy to demonstrate the apparent psychometric function of the hemispheres is that the dominant hemisphere functions as a warehouse and the non-dominant hemisphere performs the executive functions of the front office. The warehouse may contain all the necessary materials; but, if business is to function smoothly, directions must be furnished from the front office. Of course, this analogy is crude as any broad generalization, but, with this reservation, may express the principal differentiation fairly well.

TABLE 1. FIRST ORDER MEAN V SCORE DIFFERENCES BETWEEN LATERALITY GROUPS ON THE ELEVEN WECHSLER-BELLEVUE VARIABLES EMPLOYED.

Variable	NDH		DH		Mean Diff.	t	Pt(t)
	Mean	SD	Mean	SD			
Info	8.87	2.42	11.40	3.09	−2.53	2.414	<.05
Comp	5.40	2.89	7.80	3.58	−2.40	1.950	——
Arith	11.53	3.58	15.80	4.82	−4.27	2.660	<.02
Sim	11.00	3.69	14.20	3.10	−3.20	2.485	<.02
Voc	6.87	3.54	9.93	2.98	−3.06	2.481	<.02
PA	11.27	3.60	10.87	3.20	.40	——	——
PC	11.93	3.38	10.87	3.81	1.06	——	——
BD	11.13	2.42	8.60	3.09	2.53	2.413	<.05
OA	8.47	4.00	6.73	3.84	1.74	1.171	——
P5	11.80	2.43	10.80	2.34	1.00	1.109	——
V6	10.13	2.13	12.73	2.24	−2.60	3.152	<.01

TABLE 2. SECOND ORDER MEAN V SCORE DIFFERENCES BETWEEN LATERALITY GROUPS ON SUMS OF PAIRS OF WECHSLER-BELLEVUE VARIABLES.

Sum	NDH		DH		Mean Diff.	t	P(t)
	Mean	SD	Mean	SD			
Info + Comp	13.60	2.55	19.20	5.74	-5.60	3.335	<.01
Info + Arith	20.40	4.33	27.20	5.15	-6.80	3.778	<.01
Info + Sim	19.87	3.83	25.60	4.85	-5.73	3.470	<.01
Info + Voc	15.73	5.11	21.33	5.50	-5.60	2.792	<.01
Info + V6	19.00	3.63	24.07	4.40	-5.07	3.321	<.01
Comp + Arith	16.93	4.12	23.47	6.57	-6.54	3.279	<.01
Comp + Sim	16.40	5.12	22.00	5.42	-5.60	2.811	<.01
Comp + Voc	12.27	4.12	17.73	5.65	-5.46	2.924	<.01
Comp + V6	15.53	3.20	19.87	6.00	-4.34	2.385	<.05
Arith + Sim	22.53	5.34	30.00	6.97	-7.47	3.182	<.01
Arith + Voc	18.40	5.76	25.87	5.72	-7.47	3.441	<.01
Arith + V6	21.67	4.97	28.53	6.45	-6.86	3.154	<.01
Sim + Voc	17.87	5.13	24.13	4.70	-6.26	3.367	<.01
Sim + V6	21.07	4.12	26.93	5.03	-5.86	3.376	<.01
Voc + P5	18.00	2.99	20.73	2.72	-2.73	2.529	<.02
Voc + V6	17.00	4.91	22.67	4.54	-5.67	3.170	<.01

TABLE 3. SECOND ORDER MEAN V SCORE DIFFERENCES BETWEEN LATERALITY GROUPS ON DIFFERENCES BETWEEN PAIRS OF WECHSLER-BELLEVUE VARIABLES.

Difference	NDH		DH		Mean Diff.	t	P(t)
	Mean	SD	Mean	SD			
Comp – BD	–5.73	3.64	–0.80	5.74	–4.93	2.715	<.02
Arith – PA	+0.27	5.60	+4.93	5.65	–4.66	2.194	<.05
Arith – PC	–0.40	5.67	+4.93	6.77	–5.33	2.261	<.05
Arith – BD	+0.40	4.33	+7.20	5.15	–6.80	3.779	<.01
Arith – OA	+3.07	5.00	+9.07	5.46	–6.00	3.033	<.01
Arith – P5	–0.27	4.42	+5.00	4.90	–5.27	2.987	<.01
Sim-PA	–0.27	3.92	+3.33	4.54	–3.60	2.244	<.05
Sim – PC	–0.93	5.31	+3.47	4.53	–4.40	2.359	<.05
Sim – BD	–0.13	3.83	+5.60	4.85	–5.73	3.470	<.01
Sim – OA	+2.53	5.34	+6.80	3.99	–4.27	2.396	<.05
Sim – P5	–0.80	3.87	+4.07	4.25	–4.87	3.169	<.01
Voc – BD	–4.27	5.10	+1.33	5.50	–5.60	2.793	<.01
Voc – OA	–1.60	6.58	+3.20	4.89	–4.80	2.191	<.05
Voc – P5	–4.93	5.20	–0.87	4.62	–4.06	2.189	<.05
PC – V6	+1.73	4.17	–1.87	4.32	+3.60	2.243	<.05
BD – V6	+1.33	4.01	–4.13	4.47	+5.46	3.405	<.01
OA – V6	–0.20	4.85	–5.33	3.84	+5.13	3.105	<.01
P5 – V6	+1.67	3.72	–1.93	3.47	+3.60	2.647	<.02

SUMMARY

A statistical analysis of the Wechsler-Bellevue subtest differences between a group of dominant hemisphere brain-damaged patients and a group of non-dominant hemisphere brain-damaged patients is presented.

A short discussion suggests that the functions of the non-dominant and dominant cerebral hemispheres may be compared to those of a front office (executive) and a warehouse (storage), respectively. The patient with dominant hemisphere damage forgets *what* to do; the patient with non-dominant hemisphere damage forgets *how* to do.

BIBLIOGRAPHY AND REFERENCES

1. ANDERSEN, A. LLOYD. The effect of laterality localization of brain damage on Wechsler-Bellevue indices of deterioration. *J. clin. Psychol.*, 1950, 2, 191-194.
2. JASTAK, JOSEPH. Problems of psychometric scatter analysis. *Psychol. Bull.*, 1949, 3, 177-197.
3. RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic Psychological Testing*, Vol. I. Chicago: Year Book Publishers, 1945.
4. WECHSLER, D. *The Measurement of Intelligence.* Baltimore: Williams and Wilkins Co., 1939. 2nd ed. 1941; 3rd ed. 1944.
5. WECHSLER, D. *The Wechsler-Bellevue Intelligence Scale, Form II.* New York: Psychological Corporation, 1946.

THE STEREOGNOSTIC CAPACITY OF BRAIN INJURED AS COMPARED WITH FAMILIAL MENTAL DEFECTIVES

WILLIAM SLOAN
GERARD J. BENSBERG

Lincoln State School and Colony

Introduction

This problem arose out of the work done by Benton and Schultz (1) on tactual form perception in preschool children. The authors defined stereognosis as "the ability to recognize an object solely by means of the mechanism of active touch, without the use of vision." The underlying purpose of the above experiment was to develop a test utilizing tactual perception in the diagnosis of central nervous system lesions. Benton and Schultz review the neurological literature and reach the conclusion that astereognosis, i.e., impairment of stereognostic capacity, usually, but not always, indicates a central nervous system lesion. Since a controlled experiment had not been reported which investigated the effects of various variables (upon stereognostic ability) such as chronological age and type of objects to be identified, they investigated these variables. The test which they used was divided into two forms with eight objects to be identified in each form. The two forms had been found to be of equal difficulty. Their procedure was as follows: The test objects were first shown to the child and he identified them verbally. The subject's favored hand was inserted in a box with a cloth screen hiding his hand from view. The objects then were placed in his hand, one at a time, and the subject was asked to feel of it and tell the examiner the name of the object.

Benton and Schultz report no differences in score between boys and girls. The test scores or stereognostic capacity shows an increase as the chronological age increases. The authors reach the conclusion that if a child scores significantly lower on the test than others in his age group, he or she is considered defective in stereognostic capacity and a lesion in the central nervous system is suggested.

In the above experiment, the authors used only "normal" children and did not utilize any clinically diagnosed brain injured children nor did any of the children reveal overt neurological impairment. Because of this reason the present experiment was performed utilizing brain damaged subjects who, if their hypothesis is correct, should reveal astereognosis. The purpose of this experiment is to determine whether or not three clinical groups of mental defectives, familial, brain injured, and undifferentiated, perform differently on a test measuring stereognostic ability and to determine the effects of mental age and chronological age upon this ability.

Procedure

All patients used in this study were drawn from the files of patients at the Lincoln State School and Colony. Three types of subjects were compared in this study. A familial or endogenous mental defective is one who has at least one parent proven feebleminded, and in most cases, one or more siblings are feebleminded. A subject was chosen as brain injured if a diagnosis of Little's Disease, post-infection (encephalitis and congenital syphilis), or birth injury was made, in the absence of presumed hereditary factors. This diagnosis was made on the basis of an actual history of injury to the brain which is confirmed by a neurological examination by the medical staff. The third group was the undifferentiated mental defective which is a diagnosis given to defectives with unknown or unconfirmed etiology.

These three groups were matched for mental and chronological age in triplets. That is, a brain injured subject was matched with a familial and undifferentiated subject of the same chronological and mental age. An analysis of variance performed to determine whether or not the subjects were accurately matched showed no significant differences in C. A. or M. A. between the three groups. The population used

in this study consisted of 216 mental defectives, 144 being males and 72 being females. In chronological age they ranged from 10-10 to 53-1 with a mean of 28-10. In mental age they ranged from 3-1 to 8-1 with a mean of 5-9. In a pilot study it was found that subjects below three years mental age could not visually identify the objects used in this experiment.

The objects used in this experiment, which were the same objects used by Benton and Schultz, are as follows: a rubber ball, table fork, pencil, toy ring, toy chair, sun glasses, Yale type key, drinking cup, teaspoon, comb, table knife, cardboard box, stiff covered book, water glass, toy table, and a penny. A subject was brought into the testing room and his preferred hand was determined by having him draw pictures. A subject was not used if he could not name all the articles presented to him visually before the actual experiment. The subject placed his hand in the covered box and each of the 16 objects was placed in his hand for him to identify.

Results

Since it was found that the subjects performed equally well on both forms of the test, the sixteen items were grouped for a single score. This score is the number of objects correctly identified by touch.

An analysis of variance was performed to test for any significant difference obtained between the three clinical classifications and for any significant differences between the different mental ages of the subjects. Table 1 shows the summary of this analysis of variance. The F test for clinical classification was not significant indicating that the three types of mental defectives, familial, brain injured, and undifferentiated do not perform differently from each other. Since individuals within M. A. groups was not significant it was combined with Remainder, so that the Within was used as the error term in all subsequent F tests. The F test for mental age levels is significant which shows that as the mental level changes the scores made by the three groups also change. By an examination of the scores one can see that the brighter the subject the higher his score.

Table I. Summary of Analysis of Variance.

Source	df	Sum Squares	Mean Squares	F Test
Total	215	355.00		
Between	(11)	(67.66)		
Clinical Class	2	5.83	2.915	2.070
Mental Age	3	52.33	17.443	12.388**
CC x MA	6	9.50	1.583	1.124
Within	(204)	(287.34)	(1.408)	
Ind/MA	68	108.67	1.598	1.217
Remainder	136	178.67	1.313	

**Significant at less than the 1% level of confidence.

In order to test the effect of chronological age holding mental age constant, a "t" test was performed between 30 matched subjects. Fifteen randomly selected defectives with a chronological age of less than 15 years were matched for mental age with fifteen randomly selected defectives over fifteen years. The t of .403 was not significant revealing no effect from chronological age when mental age is held constant.

The correlation between mental age and the score made on the stereognosis test by the brain injured was .32. The correlation between mental age and score obtained by the familial subjects was .42. The correlation between mental age and score made by the undifferentiated subjects was .36. All of these Pearson correlations are significant at less than the 1% level of confidence showing that as mental age increases the score made by each group also tends to increase.

The intercorrelations between the score made by the three groups are as follows: brain injured and familial .12, familial and undifferentiated .18, and brain injured and undifferentiated .23. Only the last correlation borders on significance at the 5% level. These correlations reveal that no consistent relationship exists between the scores made by one group and those made by another.

Partial correlations were computed to find the correlation between mental age and score holding clinical classification constant. These partial correlations were from .05 to .09 points lower than the intercorrelations between mental age and stereognosis score for the three groups. This decrease does not decrease the correlations to nonsignificance.

Discussion

It was felt that if astereognosis, as measured by this test, reveals central nervous system lesions, it should be more prevalent in the brain injured group than in the familial group. This hypothesis, which was suggested by Benton and Schultz, was not borne out. The only essential difference in methods in this experiment and those followed by Benton and Schultz is the fact that our subjects were considerably older. An hypothesis which might be advanced to account for no astereognosis in our B-I group is equipotentiality (2). Time and experience may have operated to improve the stereognostic ability of the brain injured group.

A more plausible explanation for the failure of this test to differentiate adequately between brain injured defectives and other defectives is that no real differences exist in stereognostic ability. This finding does not refute the presumptive evidence for a central nervous system lesion existing when astereognosis is present, but it does indicate that astereognosis does not exist in all brain injured patients.

The finding of an increase in stereognostic ability with an increase in mental age is compatible with the finding of Benton and Schultz showing that it is a developmental type of process. The present authors feel that at least part of this is due to an increasing ability to focus one's attention on the object and to delay giving a response until one is sure of an answer, an attribute which is related to mental age.

Summary

1. A test utilizing the ability to recognize objects through the tactual senses was administered to 216 mental defectives. Three clinical types, familial, brain injured, and undifferentiated, were matched to find differences, if any, between the groups.

2. An increase in stereognostic capacity is noted with an increase in mental age.

3. No differences were found between the three clinical groups, which casts doubt upon the presence of a stereognosis defect, as measured by this test, in the majority of our brain injured patients.

Bibliography

1. Benton, A. L. and Schultz, L. M. Observations on tactual form perception (stereognosis) in preschool children. *J. clin. Psychol.* 1949, 5, 359-364.

2. Lashley, K. S. Basic neural mechanisms in behavior. *Psychol. Rev.*, 1937, 1-24.

A COMPARISON OF RECALL AND RECOGNITION TYPES OF MEASUREMENT ON VERBAL ITEMS, AND THEIR IMPLICATIONS FOR DETERIORATION TESTING*

E. E. GLIK

Fitzsimons General Hospital

Problem

The more widely used tests for determining intellectual deterioration employ a comparison of scores on abilities which presumably are relatively insensitive to mental deterioration with other abilities in which performance is believed to fall off rather markedly under the effects of deteriorative processes. Most of the instruments employing this method of comparison use as indices of present functioning such abilities as manipulative performance, memory, or abstraction. In almost all of these comparison type instruments, the criteria of estimated past or premorbid functioning are scored abilities in vocabulary or other verbal abilities which are believed to have remained relatively constant following their peak. These instruments, including Wechsler's Memory Scale, the Hunt Minnesota Test for Organic Brain Damage, the Babcock, the Shipley-Hartford Retreat Scale and Wechsler's Mental Deterioration Index, have all been subject to criticism, both in the literature and among clinicians using the instruments. One reason for doubting the validity of deterioration scores derived from such tests is their common underlying assumption that the subject's verbal skills are premorbidly of the same magnitude as his skill in visual motor performance, abstraction, and other abilities. We know from the many inter-test scatter studies that individual differences in verbal as compared with other abilities are wide.

Juckem and Wold[1] found that the Hunt Minnesota Test for Organic Brain Damage was practically useless at the upper levels of vocabulary. The subjects of their study—although clinically without organicity—consistently received spuriously high deterioration scores simply because their verbal abilities were noticeably in excess of their abilities in other areas measured by the test. It seems reasonable to suppose that other instruments designed to detect deterioration by the method of comparing vocabulary to other skills would likewise prove ineffective in highly verbal subjects unless the clinician knew something about the patient's differential between verbal and other abilities. On the other hand, one would also reasonably suppose that subjects whose vocabulary was never on a level as high as other mental abilities, would tend to receive scores which possibly fail to pick out all but the most deteriorated of organic cases. In other words, the inability to estimate a patient's premorbid differential between verbal and other abilities has caused these types of mental deterioration indices, memory quotients, or deterioration scores to be valid only for those patients whose verbal ability never varied greatly from other abilities.

The purpose of this study is to test out a clinical method designed to improve the validity of deterioration appraisal. It aims to do this by providing a more accurate estimate of premorbid or "hold" ability in the verbal area, and to demonstrate a means of detecting the effects of memory dysfunction upon verbal ability itself. The method devised is based on a fact observed by Luh[2] in 1922 and now familiar in human learning and memory; namely, that retention of learned verbal material is greater when measured by recognition than when measured by direct recall. It may be assumed that in many instances, particularly where mental deterioration has been incurred, test scores in the verbal area fail to accredit the patient with material that he may have known at one time, but can now recall only by recognition. This

*A paper read before the Annual Meeting of the Rocky Mountain Psychological Association, Fort Collins, Colo., May 13, 1950.

assumption is attested to by the frequent instances in which deteriorated patients are observed to reply to questions of fact with such statements as "I should know that. It's on the tip of my tongue but I just can't think of it."

The Shipley-Hartford is the only deterioration instrument that employs a recognition type of vocabulary through the use of multiple choices offered for each word item. In this respect the Shipley-Hartford may provide a better estimate of premorbid verbal ability than other tests mentioned. The presentation of the word stimuli along with a choice of correct and incorrect responses would seem to increase the validity of deterioration scores in predominantly non-verbal subjects. Neither the Shipley-Hartford nor the other instruments mentioned however, employ a comparison of the subject's recall and recognition vocabularies in an attempt to detect deterioration of verbal ability itself, and thereby furnish a more direct means of appraising deterioration in predominantly verbal subjects.

The present study is an attempt to demonstrate the effects of deterioration on the ability of subjects to recall known verbal material, and to suggest the clinical usefulness of employing a recognition type of verbal measurement in conjunction with deterioration instruments now in use. The experiment consisted of presenting to organic and non-organic subjects of various ages, multiple choice lists with the Vocabulary and Information Subtests of the Wechsler-Bellevue Intelligence Scale, Form I. The number of items on the multiple choice list correctly recognized by each patient was compared to the number of items correctly answered by him in the orthodox method of direct recall. If deterioration of verbal ability itself is measureable, and if a recognition type of measurement furnishes a more accurate premorbid estimate of verbal ability, then the disparity between recall and recognition measurements of verbal ability should be found to be greater in older and deteriorated individuals.

Subjects

Multiple choices were presented in a manner to be described to fifty-six adults hospitalized as neuropsychiatric patients in an Army general hospital. In order to attain some measure of control of deterioration factors due to age alone, the subjects were grouped according to their ages, which ranged from 19 to 62 years. One age group consisted of patients below age 45 and the other age group contained patients above 45 years of age. Approximately half of the members of each age group had incurred some clinically diagnosed deterioration and were compared to patients of similar age who did not exhibit evidence of deterioration. The criterion of deterioration in each case was the final diagnosis made by the staff of the neuropsychiatric service, which diagnosis in turn was based upon clinical and laboratory findings. The data obtained in this study were in no way used in arriving at these diagnoses.

Table 1 shows the statistical description of the sample population. The slight difference in the age distributions between the deteriorated and undeteriorated group for each of the two age categories are not significant, and the difference between the mean age of any young group— deteriorated or not—and the age mean of any old group is at least as large as two standard deviations taken respectively from each mean. Thus we do have two dissimilar populations represented as far as age is concerned.

A rather crude attempt was made to control intelligence as a variable possibly affecting verbal retention. Only patients whose Full Scale IQ score fell within the normal range of 90-110 were eligible as subjects. Although such a restriction probably means that the deteriorated group possess a slightly higher premorbid intellectual level than the present intellectual level of the non-deteriorated group, this was considered to be the only feasible method of obtaining some degree of control of intelligence.

The sample is unfortunately small because of the nature of the patient population in an Army neuro-psychiatric service. It was particularly difficult to find patients who were mentally deteriorated. Such individuals are usually screened out

TABLE 1. STATISTICAL DISTRIBUTION OF CASES STUDIED IN NOT DETERIORATED AND DETERIORATED GROUPS.

	YOUNG		OLD	
NOT DETERIORATED	N	Age	N	Age
Obsessive-Compulsive	6		4	
Hysteria	5		6	
Anxiety	3	M 34	2	M 56
Reactive Depression	1	R 19–42		R 49–62
Chronic Alcoholism	1	SD 7.23		SD 4.0
Psychopathic Personality	1			
Involutional Melancholia			2	
N	17		14	
DETERIORATED	N	Age	N	Age
Post Traumatic	4	M 32		M 56
Epileptic	3	R 24–43		R 52–62
Alcoholic	2	SD 6.75	4	SD 2.75
Huntington Chorea	1			
Multiple Sclerosis	1			
Hebephrenic Schizophrenia	2			
General Paresis			4	
C. A. S.			4	
N	13		12	

at Army induction centers and those who incur deterioration while in service are not long in the Army's custody. It is hoped that other investigators will continue similar study of this problem in state mental institutions where there is a greater need for its application, and where a larger, better matched, and more adequate sampling of deteriorated cases may be found than is described here.

PROCEDURE

The procedure of the investigation consisted of presenting the multiple choices shown below to the subjects immediately after they had been administered the Information and Vocabulary tests in the conventional manner. The number of items on each of the tests correctly answered in the conventional or direct recall manner was computed for each subject as a percentage of the number of items correctly answered later on the multiple choice list. In other words, if the patient's measured vocabulary was totally elicitable by direct recall he received a percentage score of 100. If a portion of his vocabulary could be recalled only by the use of the multiple choices, his score would be that much less than 100%. Lower scores, it is presumed, would tend to indicate memory impairment. The construction and the administration of the modification were designed to keep chance selection factors operating at a minimum. For example, the subjects were instructed not to make guesses but to select a multiple choice only when they thought it meant the same or nearly the same as the stimulus word. To further insure against guessing an attempt was made to include in each response cluster a word which—if the subject were naive of the stimulus word's meaning—would seem the most correct. Thus the word "corpse" is one of the possible responses to the question, "What is an Habeas Corpus?", and the word "limbs" is offered as a possible response to the meaning of the Vocabulary item, "Armory". Both of the February holidays appear in the response choices to the Information item requesting the date of Washington's birthday. The choice lists have already been subjected to revision and will probably be revised again in order to further limit factors other than recognition and recall from determining a subject's responses. The present lists are the most revised and the ones used in the current research.

Vocabulary

1. Apple—basket, fruit, sin, fish
2. Donkey—hay, large, animal, cars
3. Join—connect, ring, borrow, say
4. Diamond—drink, woman, jewel, pair
5. Nuisance—bird, noise, neck, pest
6. Fur—skin, floor, trap, bear
7. Cushion—start, pillow, profane, nice
8. Shilling—world, splinter, coin, planting
9. Gamble—penny, bet, reach, run
10. Bacon—cook, money, try, meat
11. Nail—spike, hang, cat, kill
12. Cedar—snake, holiday, tree, taste
13. Tint—color, camp, refrain, jump
14. Armory—arsenal, battle, limbs, set
15. Fable—weak, dumb, tale, crazy
16. Brim—indicate, pool, edge, down
17. Guillotine—toy, behead, alcohol, feat
18. Plural—water, sharp, up, many
19. Seclude—hide, refuse, death, joy
20. Nitroglycerine—holy, ink, explosive, country
21. Stanza—vine, declaim, perch, verse
22. Microscope—heavens, instrument, long, measure
23. Vesper—hymn, coat, soften, west
24. Belfry—order, steeple, harden, out
25. Recede—tower, point, withdraw, win
26. Affliction—conclusion, please, office, malady
27. Pewter—metal, tea, distant, fleece
28. Ballast—grief, art, weight, fry
29. Catacomb—brush, groom, underground, horse
30. Spangle—decoration, wave, sing, hit
31. Espionage—complete, spying, hang, joy
32. Imminent—incidental, poor, impending, lost
33. Mantis—battle, plant, host, insect
34. Hara-Kiri—suicide, news, plane, ride
35. Chattel—herd, beef, property, miser
36. Dilatory—meritorious, leading, delaying, fierce
37. Amanuensis—machine, secretary, medicine, cure
38. Proselyte—convert, predict, toast, feast
39. Moiety—largeness, pity, certainty, half
40. Aseptic—cure, harmful, sterile, latent
41. Flout—dance, snub, jeer, whip
42. Traduce—conclude, trace, slander, strengthen

Information

1. Lincoln, Roosevelt, Washington, Hoover
2. Instrument, bomb, container, organization
3. Earth, tires, trees, natives
4. France, England, Italy, Boston
5. Four, six, eight, two
6. 48, 4-1/2, 52, 365
7. London, Paris, Naples, Rome
8. Shanghai, Berlin, Rome, Tokyo
9. 5′0, 6′0, 4′5″, 5′5″
10. Wiley Post, Edison, Wright Bros., Marconi
11. England, South Africa, South America, Australia
12. 3,000 miles, 5,000 miles, 1,000 miles, 10,000 miles
13. Cleans body, pumps blood, controls body temperature, breaks
14. Othello, Milton, Longfellow, Shakespeare
15. 130,000 - 200,000,000 - 1,000,000,000 - 10,000,000
16. Jan. 1 - Feb. 12 - July 4 - Feb. 22
17. Byrd, Columbus, Erickson, Peary
18. America, Palestine, Russia, Africa
19. Shakespeare, Hawthorne, Mark Twain, Ring Lardner
20. World's capital, UN headquarters, Pope's home, European University
21. Secret organization, poem, Mohammedan Bible, myth
22. Shakespeare, Goethe, Gounod, Bizet
23. Legal writ, corpse, murder charge, crime
24. Study of insects, study of races, study of breeding, study of words
25. Protocols of Zion, catastrophe (myth), Writings of Old Testament, Building in Ancient Greece

TABLE 2. COMPARISON OF VOCABULARY AND INFORMATION SCORES OF YOUNG AND OLD, DETERIORATED AND NON-DETERIORATED GROUPS

YOUNG	Range	Mean %	SD
Non-deteriorated—Vocab.	90.5–100	96	2.25
Deteriorated —Vocab.	79 – 96	88	2.50
Non-deteriorated—Info.	93 –100	96	2.93
Deteriorated —Info.	54.5– 86	70	10.50
OLD			
Non-deteriorated—Vocab.	82 –100	91.5	5.35
Deteriorated —Vocab.	91 – 96	89.5	5.79
Non-deteriorated—Info.	83 –100	93	5.85
Deteriorated —Info.	75 – 91	85.5	6.32

The data for all groups is presented in tabular form in Fig. 2. The "*t*" test of significance is applied to these comparisons with the following results. The probability of the difference in mean vocabulary recall scores between the deteriorated and undeteriorated young groups being due to chance alone is less than 1%. A comparison of the means on Information recall for the young deteriorated and non-deteriorated subjects cannot be made because of the large difference in variance observed. The discreteness of the two distributions, however, indicates a significant difference. The "*t*" test applied to the two vocabulary means of the old group indicate that the Null Hypothesis concerning differences between the deteriorated and non-deteriorated subjects can be rejected only between the 20% and 30% levels of confidence. For this age group, however, the similar Null Hypothesis concerning the differences in Information recall can be rejected at better than the 1% level.

Testing now the significance of differences presumably due to age alone, we find that those of the non-deteriorated group are the only ones in which age differences appear in the expected direction. The difference between the mean Vocabulary percentage recall scores of the young non-deteriorated group and the old non-deteriorated group are significant at the 1% level. Comparing these same groups in their Information percentage recall scores yields a "*t*" which is significant between the 10% and 20% levels of confidence.

The results of similar comparisons between the old and young of the deteriorated subjects yields results contrary to expectation. Thus, there is no appreciable difference between the young and old deteriorated subjects in their percentage recall scores in Vocabulary. The difference of 1.5% could have occurred almost purely by chance. Comparing the young and old deteriorated subjects in their Information recall reveals that the older subjects depended less on the multiple choices than did the younger subjects. This apparent contradiction is probably explicable if one reasons that at least some of the deteriorated subjects in the young group were deteriorating much more rapidly than those in the old group; and that these younger deteriorated subjects would probably not have even been testable at an age corresponding to that of the old group, if indeed they could be expected to survive long enough to reach such an age. This explanation would seem to be true particularly for the cases of Huntington's Chorea and Multiple Sclerosis, whose percentage recall scores in Information were the lowest for their group.

It is to be observed that not all failures to recall known verbal material are indicative of organic impairment. Very few of even the young and non-deteriorated subjects attained recall percentage scores of 100. The element of anxiety in effecting transient difficulty in verbal recall on these very same tests has recently been reviewed by Rashkis and Welsh[3]. In this study the cases diagnosed as anxiety states, reactive depressions, and involutional melancholias received the lowest percentage recall scores in the non-deteriorated groups.

Conclusions

Inferences drawn from the foregoing results should be made with some caution because of the smallness of the sample and because certain factors, such as extent of deterioration and premorbid intelligence, were not properly matched or controlled. It has been demonstrated that the disparity between recall and recognition measures of Vocabulary and Information is greater in individuals who have incurred intellectual deterioration. This finding supports the earlier assumption that a recognition type of measurement provides a more accurate premorbid estimate of verbal ability. It also warrants the further experimentation with the recognition type of measurement in conjunction with deterioration testing where the clinician wishes an accurate premorbid estimate of intellectual ability, or where it is desired to appraise deterioration without risking the assumption of premorbid equality of many varied abilities.

Although the purpose of this study is not to establish norms by which to predict the existence of deterioration, the clinician wishing to use these findings and these multiple choices with the necessary reservations, might be suspicious of deterioration when patients under 40 years of age can recall less than 90% of what they can recognize correctly on Wechsler's Verbal items, and in patients over 50 years of age who can recall less than 80% of the material correctly recognized. These approximations would be slightly less for Information than for Vocabulary. Wechsler's observation that Vocabulary holds up better with deterioration than Information is also observed in this study.

References

1. Juckem, H. and Wold, J. A. A Study of the Hunt Minnesota Test for Organic Brain Damage at the Upper Levels of Vocabulary. *J. Consult. Psych.*, 1948, 12:53-57.
2. Luh, C. W. The Conditions of Retention. *Psychol. Monograph*, 1922, 31, No. 142.
3. Rashkis, H. A. and Welsh, G. S. Detection of Anxiety by Use of the Wechsler Scale. *J. Clin. Psychol.*, 1946, 2, 354-357.

A COMPARISON OF THE STIMULUS VALUE OF SZONDI'S PICTURES WITH THOSE OF NORMAL AMERICANS

WILSON H. GUERTIN

*Michigan State College**

Introduction

The Szondi pictures are of such an alien character that many workers must have experienced a bit of a shock upon seeing them for the first time. These pictures from an European mental hospital population are being applied to Americans assuming that the meaning for the subjects will be the same as that for the European subjects, upon whom the interpretation of picture selections was based. One wonders if certain advantages might not obtain through using photographs of American hospital cases', and might even question the value of using the mentally ill for stimulus pictures. The answer to both questions would be empirical. It should not be difficult to devise a set of pictures of normal Americans which could be compared with the Szondi set. But what would be the measure of their respective diagnostic values? There probably are many possibilities, one of which is to be considered in detail in this paper.

We might look to item analysis for such a measure, since this mathematico-deductive technique does not require many empirical assumptions. Of particular interest is the concept of *discriminating power* (3, p.426ff). A test is said to have high discriminating power if it differentiates well between subjects who have differences in the quality being measured. i.e. with comparable reliabilities, the discriminating power of an item is a function of the range of scores obtained with the item when a heterogeneous group of subjects is employed.

This consideration of the importance of a range of scores for a test item suggests a measure of the relative value of projective stimuli without reference to the nature of the qualities being measured. If unreliability is held equal, the stimulus pictures that produce the greatest variability in preference ratings would be of the greatest value in reflecting individual differences in the subjects.

Purpose

This paper reports the results of a comparison of the Szondi pictures with a set of 48 American non-mentally ill individuals. The chief question to be answered is: Do the Szondi pictures provide better stimuli for evaluating individual differences than the control set? Incidental information should disclose which of the Szondi pictures are the better ones from this point of view.

Procedure

Thirty-two, white undergraduates served as subjects, volunteering from an introductory course in general psychology. Superficially, the group is fairly homogeneous, but wide personality differences are presumed to exist. An equal number of each sex was used to provide a considerable range of personality variables.

The experimental condition utilized the Szondi pictures while the control employed a specially selected group of 48 photographs of non-hospitalized individuals which had appeared in various recent American magazines. The photographs were of scientists, models, statesmen and a few others. All presumably were American citizens. Since these pictures were matched with the individual Szondi photographs for size, background, sex, and approximate age; the chief difference was that they were not foreign or mentally ill. Since pictures in magazines are generally smiling some selection had to be employed to prevent this characteristic from being too prominent in the control set.

*Now at Beatty Memorial Hospital, Westville, Indiana.

The experimental design is such as to reveal the variability of preference from individual to individual for the pictures in each of the two sets. The subjects were requested to rank the pictures from most to least preferred. They also ranked the control set to provide a basis of comparison for the magnitudes of the variability of the picture preference values in the two sets. Counting procedures, applicable to rank-order scales, were employed to indicate the average preference value and the dispersion about it. Median preference ranks and semi-interquartile ranges are the statistics used. However, another variability measure was introduced for two reasons. First, the range is a rather unreliable measure, and second, the measure introduced has practical importance in indicating the discriminating power of a given picture. This measure was devised for this study and is called *equipotentiality*. It is merely the probability expressed in percentage that a subject has for selecting a picture as like or dislike, depending upon his personality. If a picture is selected 30 percent of the times as liked and zero percent as disliked, a subject is not likely to select it as disliked. The equipotentiality would be zero, the smaller of these two percentages.

Results

The distributions of the various ranks assigned to a given picture were of various forms. They were most frequently rectilinear with skewing when the median was near the end of the scale. Desirable bi-modal distributions were not in evidence except for one picture in the control set.

Table 1 presents the two measures of variability in ranks assigned to the individual pictures of both sets. A fair correspondence in values is apparent. The attenuating effect of the scale ends is kept equal in this comparison by arranging the individual pictures of each set according to their median preference-value ranks. When corresponding values of the semi-interquartile ranges for the pictures of the two sets are compared, it is seen that the Szondi values are higher in 22 cases while the control values are higher 24 times. A similar comparison of equipotentialities indicates that the Szondi was higher 16 times as compared with 19 times for the control. Similarly small differences in frequencies are obtained through the use of cutting-scores in a comparison of the distributions of the magnitudes of these variability measurements.

Since the variability of the rankings appears to be comparable for the two sets, it will be necessary to demonstrate that neither set has a greater unreliability. This was done by obtaining retests on six subjects, three of each sex, after an interval of one week. Rank-difference correlation coefficients expressing the relationship between the subject's ranking and reranking of the pictures were calculated. They were corrected to Pearson "r" 's and combined through "z" transformations[4, p. 213] to provide single estimates of the reliability coefficients for each set. They were found to be .745 and .747 for the Szondi and control respectively. The SE of the difference between them is only .052.

Further analysis of these retest data seemed desirable to determine whether the Szondi pictures with higher variability in the rankings reflect better discriminating power or merely greater unreliability. The comparison made was between the 24 pictures showing the higher variability and the 24 with the lower. The sum of the test-retest discrepancies in rank positions for these two groups of pictures failed to reveal any difference in the reliabilty of them.

Discussion

The results of this study fail to affirm the unique value of Szondi's stimulus pictures. The pictures of non-mentally ill Americans used as a control series showed equally high variability for the rankings, presumably related to individual differences. These individual differences must remain undefined but they would probably lie within the realm of personality since it is difficult to see how other variables

TABLE 1. VARIABILITY IN THE RANKING OF THE PICTURES OF BOTH SETS, IN ORDER OF PREFERENCE RANKS.

Preference (*high to low*)	*Szondi Pictures*	*Szondi*		*Control*	
		S. I. Quart.	*Equipot.**	*S. I. Quart.*	*Equipot.**
1.	Vh	1.9	3.1	1.3	3.1
2.	IVs	1.8	0.0	2.2	3.1
3.	IVp	5.6	3.1	2.9	6.2
4.	IIm	3.3	3.1	4.3	6.2
5.	IVk	1.0	3.1	3.6	6.2
6.	Vm	4.8	3.1	4.6	9.4
7.	Vp	4.2	6.2	4.0	3.1
8.	Im	1.9	3.1	5.0	9.4
9.	IVe	3.2	6.2	4.1	0.0
10.	VIh	6.7	15.6	1.7	6.2
11.	IIp	5.7	9.4	3.7	9.4
12.	IIIm	6.0	15.6	4.9	9.4
13.	IVh	3.5	9.4	1.8	12.5
14.	Ie	4.2	15.6	5.0	15.6
15.	VIk	7.1	21.9	5.1	15.6
16.	IIIp	6.1	15.6	3.2	6.2
17.	Vhy	4.6	12.5	7.1	28.1
18.	IId	5.5	12.5	7.4	21.9
19.	VIm	3.4	15.6	3.9	12.5
20.	VIe	7.0	21.9	6.0	18.8
21.	IIIe	4.9	21.9	6.4	21.9
22.	Vk	3.4	15.6	5.2	18.8
23.	Vd	4.9	12.5	5.0	3.1
24.	IIk	2.1	9.4	6.8	9.4
25.	IIs	11.2	31.3	6.5	18.8
26.	IVhy	5.5	15.6	5.5	0.0
27.	Ih	3.4	15.6	5.1	9.4
28.	Ihy	5.2	18.8	6.3	21.9
29.	VIhy	4.5	9.4	7.4	25.0
30.	IIId	4.2	0.0	4.3	15.6
31.	IVm	2.3	0.0	14.0	34.4
32.	Id	3.4	9.4	3.4	9.4
33.	Vs	6.0	15.6	4.4	6.2
34.	Ve	5.6	15.6	5.4	18.8
35.	IIh	6.5	15.6	4.5	6.2
36.	IVd	5.5	12.5	9.9	28.1
37.	IIIh	6.5	6.2	4.6	6.2
38.	IIIk	5.0	6.2	4.5	12.5
39.	Ip	7.4	12.5	3.6	3.1
40.	VIp	6.1	3.1	2.7	3.1
41.	IIIs	7.4	12.5	4.2	6.2
42.	VIs	7.3	15.6	5.5	9.4
43.	IIe	4.2	0.0	6.1	9.4
44.	VId	3.2	3.1	4.0	3.1
45.	Ik	2.4	0.0	4.9	3.1
46.	Is	3.6	3.1	2.2	3.1
47.	IIhy	1.9	0.0	3.0	0.0
48.	IIIhy	1.9	6.2	3.4	6.2

*See text for the meaning of this statistic.

would influence picture preferences. Since Szondi's pictures do not show a greater dispersion of rankings it is concluded that either they do not evaluate *need-tensions* better or other personality variables in addition to Szondi's need-systems exist and to such a degree as to obscure the effect of his need-systems.

The rather wide differences in the dispersions of rankings for the various Szondi pictures suggest that some (the higher ones) should provide better diagnostic clues than the others. This greater variance observed for some of the pictures seems to be attributable to the interaction of the picture cues and personality variables rather than to greater unreliability. As in previous studies, wide differences in the Szondi stimulus pictures appear.

The control pictures are relatively unimportant except as they disclose information about the Szondi test or the general method by serving as controls. While no systematic study of the relationship between the Szondi and control sets was made with respect to preference value, the few subjects who were asked to place some of the control pictures along the scale of ranked Szondi pictures placed them all in the upper half. This behavior was concordant with verbal comments during administration.

It is interesting that the highest variability occurred for a picture in the control set. This was the picture that gave a bimodal distribution. It was of a little girl, a full-face view. The child has long and rather tangled hair, her eyes are turned sharply to the right, and her tongue protrudes partially. It is possible that the wide variability in responses to this picture arose from regarding it as attractively 'impish' or unattractively 'idiotic'.* It would appear that the first step in constructing a more efficient Szondi-type test would be to obtain such pictures as this through an analysis of the variability of a large number of potentially valuable photographs.

In conclusion, it is felt that this investigation serves to throw some doubt upon the value of Szondi stimulus pictures. The view of the writer[1, 2] is that the Szondi *method* is not being applied in a very efficient fashion by the Szondi *test*. Refinements and modifications in the technique and stimuli probably would lead to a more useful instrument for evaluating personality.

Summary and Conclusions

1. The value of the Szondi stimulus pictures is examined with respect to the use of European mental hospital cases for stimuli.

2. Thirty-two undergraduate subjects of both sexes ranked the Szondi pictures and also a set of non-mentally ill Americans in the order of their preferences.

3. A comparison of the dispersions of these rankings for the individual pictures was used as an indication of the *discriminating power* of them. The comparison failed to demonstrate the superiority of the Szondi pictures. Retest results showed that the two sets have comparable reliability.

4. The rather wide differences in Szondi stimulus-pictures suggest that some should provide more diagnosic clues than others. Those pictures with greater dispersion of rankings had retest reliability comparable to those with smaller dispersions.

5. A more careful selection of stimulus pictures probably would yield a more reliable and more valid measure of individual differences in the personality of American subjects.

References

1. Guertin, W. H. A test of a basic assumption of the Szondi test. *J. consult. Psychol.*, 1950, 14, 404-407.

2. Guertin, W. H. A consideration of factor loadings on the Szondi test. *J. clin. Psychol.*, 1950, 6, 262-266.

3. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

4. Lindquist, E. F. *Statistical analysis in educational research.* Boston: Houghton Mifflin, 1940.

*The author is indebted to Dr. Laurence F. Shaffer for this description.

THE LETTER SERIES AS A PROJECTIVE TECHNIQUE

HELEN M. WILTSHIRE

Kent State University

Introduction

This study proposes to investigate the possibility of a projective technique whereby a child reveals his personality trends through the ready, uncensored response to certain stimulus letters. The resultant response is seen as projecting from past experience and is material which has come to have meaning for the child. It is most readily available in a projective setting. The "spoken word" of any child thus becomes meaningful when interpreted in relation to other spoken words in the particular configuration. Since the child's adjustment is dependent largely on his interpersonal relationships within the family and school, it follows that these areas should be given an important place.

A list of 30 letters, arranged to duplicate M's and D's, was selected for administration. The first letter is the first letter of the child's own first name. The responses were expected to throw light on the following areas of personality: (1) Parental relationships; (2) peer relationships; and (3) concept of the self. Additional supplementary areas considered to be of interest were: (1) Play types; (2) indications of social intelligence; (3) early patterns of ego defense; and (4) the reality of school grades as a measure of efficiency at this age level.

Method

The subjects were 50 pupils, 25 of whom were 11 years of age, and 25 were 12 years of age, who volunteered to cooperate in the study. All were given the Letter Series individually, during school hours between one and three in the afternoon, and outside the classroom. The school is one of three elementary districts, serving a middle-class suburban area of 12,000. There are about 25 pupils in each grade. The pupils who volunteered for the study represented the entire sixth grade, a few 11 years old from the fifth grade, and about half of the seventh grade population. Six pupils came from a different school, but the school was similar in all respects. The pupils appeared to find the work highly interesting. Neither the children or their families were known to the writer. Kuhlmann-Anderson Group Intelligence scores were available for most of the pupils. The group tends to be average or above, as determined by this test. Incidental use only is made of the scores.

The instructions given were: "I am going to read some letters and you are to write the first word you think of that starts with this letter. Spelling will not count. There are no right or wrong answers. Just write the first word you think of as quickly as you can. If you cannot think of a word, copy the letter itself, and we will then go on with the rest." An attitude of speed was prevalent, as the children were seen in fairly rapid succession and no attempt was made to interview them or record notes on their reactions. When a pupil could not think of anything to write, however he was helped on to the next word, with full acceptance that it did not matter.

Scoring

The clinician will probably use his own individual method of constructing the protocols, depending on his frame of reference. A few suggestions may be indicated, however. The word selected for the initial stimulus letter is significant, as it may or may not be the child's own name. After this is noted, the next 4 letters are disregarded as part of the process of adjustment. The series proper consists of the last 25 letters, and the first letter of this series is M. This may or may not refer to attitudes related to the relationship with the mother. Blocking is assumed to follow a critical response. The entire response series is viewed dynamically on an interrela-

tionship basis for the total personality description. All areas may not be observed in each protocol, and interpretation is made accordingly. No scoring, as usually conceived, is attempted. The result is a descriptive paragraph of the dynamic personality characteristics revealed by the child, himself, in his selection of words. The protocols are constructed entirely from the Letter Series.

In an effort to check the validity, a rating scale was constructed, covering ten items, and the teachers were asked to record their opinions in these respects. One of the scales was filled in for each pupil. The opportunity was given for a 5 point range in rating, 1 being poor, 5 good. The specific questions included in the rating scale were: (1) How mature personally do you consider him in terms of age; (2) how mature socially do you consider him in terms of age; (3) to what extent is he a leader; (4) is he a happy child; (5) is he at all subject to fears; (6) how well does he get along with the group; (7) how well does he get along with his own sex; (8) is the relationship with the mother harmonious; (9) is the relationship with the father harmonious; (10) how does he rank in school grades.

Illustrative Cases

The following sample protocols give an indication of the material which may be elicited by the Letter Series Technique:

Subject 1: Boy, 11 years, 6th grade, IQ 95. The Letter Series follows: J, Johnny; A, Audrey; B, Boy; C, Cow; F, Floor; M, Moon; M, More; L, Lark; D, Dog; D, Dan; I, in; H, hall; S, shall; K, k; T, —; W, War; W, wall; S, short; H, hart; D, Dear; G, go; L, long; R, row; M, Mort; D, Don; C, corn; D, Dong; T, Tom; N, no; O, Oh.

Protocol: J's parental attachments are repressed. Early emotional deprivation is revealed, and may account for the hostility. He creates the impression of being well socialized, and interested in peers, also very possibly has several close friends. He shows self-awareness, but is not wholly self-accepting. Initiative is lacking, also the achievement drive. There is a lack of acceptance of his play interests and fun loving activities, and play seems to be, for him, rather a hostile, aggressive act. This affords him a mechanism of defense and means of annoyance to the parent person. He is very likely doing poor work in school, and tends to adopt the feeling that he is rejected by the teacher. His over-concern for play may arouse in her some need to remonstrate the child to try to feel more grown up, so that he has by now come to play more and more, to the discomfort of his parents and parent surrogates.

On the rating scale, the teacher indicates a rank of 3 in both personal and social maturity, and in leadership; a rank of 4 in the next 4 items; a rank of 5 in parental relationships; and a rank of 2 in school grade achievement. When later questioned about J's play interests, the teacher stated that he is athletic in his interests, generally, is greatly engrossed in play as a participant, and always wants to get outside quickly at recess.

Subject 2: Girl, 12 years, 7th grade, IQ 109. The Letter Series follows: P, pineapple; A, apple; B, ball; C, car; F, fire; M, mother; M, more; L, love; D, dog; D, dear; I, Ida; H, house; S, stairs; T, tall; K, kite; W, wagon; W, ware; S, sure; H, horse; D, dig; G, good; L, lonely; R, ray; M, mood; D, dong; C, cargo; D, darling; T, tongue; N, none; O, on.

Protocol: Phyllis is described as a steady, even tempered, conforming pupil, dependable and easily adjusted. Her feeling relationships tend to be on a tender, sensitive, rather "adoring" level, and no doubt she is capable of the idealizing and fantasying characteristic of many girls her age. Her recreational interests lean to the passive type rather than the active and there is some solitariness

in this area, suggesting she may not have as many outlets as she would like to have. Insecurities are offset by close family ties, especially to her mother. She is beginning to use fantasy as an adjustment and escape mechanism.

On the teacher's rating scale, we find a rank of 4 in personal and social maturity; a rank of 3 in leadership; 4 on the next 2 items; 3 in group adjustment; 4 in adjustment to girls; a rank of 5 in relationship with the mother, 3 with the father; and a rank of 4 in school grade achievement. The protocol was later discussed with the teacher who considered the description fairly accurate. The matter of fantasy and adjustment mechanism was left open as the teacher was not able to give this information.

Subject. 3: Girl, 11 years, 6th grade, IQ 131. The Letter Series follows: L, Light; A, Apple; B, box; C, Cat; F, face; M, Mother; M, moth; L, l; D, dog; D, duck; I, i; H, home; S, Sunday; T, Tommy; K, kiss; W, wet; W, went; S, sun; H, how; D, —; G, gone; L, —; R, run; M, mouth; D, done; C, come; D, d; T, tom; N, none; O, oval.

Protocol: Lee is closely attached to her mother, possibly is held by her on a dependency level, while the father attachment is repressed, or absent. Peer relationships are not strong, but an interest in boys is evident. She tends to be a sensitive, hesitant pupil, with insecurities in new situations, and does not feel comfortable when required to depend on herself. Early adolescent drives apparently cause some anxiety. Home is a source of security and the chief source for Lee, it would seem. One suspects she is not working up to capacity in school, but probably gets by.

On the rating scale, the teacher checks 3 for personal and social maturity; 2 for leadership; 4 for item 4; 3 for the next 3 items; 5 for parental relationships; and 3 for rank in school grade achievement. This indicates perhaps an over attachment to the parents and submissiveness in the group, an adjustment which seems satisfactory to Lee as well as to the peers. We know this situation is not conducive to optimum personal and social growth, and are not surprised to find she is not working up to capacity in school. The variability in the rating scale signifies some lack of balance in general adjustment. When later questioned about the protocol, the teacher disclosed that Lee's father is considerably older than her mother, and although he is very fond of the family and a good provider, he may not be very close to Lee on a relationship level. Lee and her younger brothers are frequently "brow-beaten" by other children on the way to school, and customarily run home to the mother for protection.

Subject 4: Boy, 12 years, 6th grade, IQ 124. The Letter Series follows: J, jet; A, all; B, bad; C, can; F, fun; M, Monday; M, Mummy; H, had; D, dandy; D, dad; I, I'll; H, had; S, sat; T, tea; K, kite; W, we; W, wed; S, see; H, has; D, do; G, got; L, land; R, rat; M, mad; D, Doody; C, candy; D, d; T, the; N, name; O, Oh.

Protocol: Relationships with the parents are close and affectionate. Peer relationships are satisfactory, and group awareness is evident. He shows self-confidence, is forceful, intelligent and out-going. There is some capacity for leadership apparent. He demonstrates self-direction, is aggressive, healthily self-indulgent. He is adopting early escape fantasies as a measure in his adjustment.

The teacher's rating scale shows a rank of 5 on 9 of the 10 items, the exception being a slight tendency to fears, with a rank of 4 on item 5. This describes a well adjusted pupil with a good personality.

Summary

The Letter Series Technique gives an indication of the type of configuration which may be elicited in this particular framework with a minimum of pupil time involved. Less than 5 minutes is required for administration. It aims to eliminate some of the difficulty of group intelligence testing by (1) the personal interview with the child, which assures some measure of control over the pupil's attitude, application and degree of cooperation; and (2) by the unstandardized answers which permit the taking into account of cultural factors in the environment. There are no right or wrong answers, but a configuration protocol as "score". The protocol may be adapted by the clinician to the interpretative level and understanding of the persons who will make use of it. A means of appraisal of the individual in terms of the group in which he is expected to function is also provided.

The Letter Series Technique is also a means by which a survey may be made and recorded of each pupil's progress from time to time. This is significant (1) for calling attention to unhealthy attitudes which may be accumulating and which at the moment, may not be upsetting to the pupil, and (2) for future research. It may be possible to observe the personality development of those individuals who will later become our leaders, our criminal population, or those who will become patients in our mental hospitals.

Bibliography

1. Lloyd, Wilma. Some Aspects of Language as Significant of Personality. *Psychol. Bull.*, 1941, 38, 747, (abstract).
2. Sanford, F. H. Speech and Personality. *Psychol. Bull.* 1942, 39, 811-841.

CLINICAL PSYCHOLOGY IN PUBLIC SCHOOLS*

DOUGLAS COURTNEY

Ohio State University

By tradition the clinical psychologist has a special role in the study and modification of human behavior. From a humble and limited beginning, determining the intellectual level of individual children, his diagnostic techniques have widened and new therapeutic areas have been opened. Despite his professional progress, the concept of individual diagnosis and individual therapy has retained dominance as his mode of practice. Operating under such a restriction, he is limited perforce to restrict further his attention to the extreme deviant. Operating, as he so often does, in a clinical setting to which children are brought from diverse environments, he seldom has available adequate descriptive statistics on the environment from which the child is drawn. Thus he is unable to offer adequate evaluations of individual behavior which will have functional significance when the child is returned to his original setting.

In Ohio (to which this report is limited), there is a growing belief that clinical psychologists have a public health function, i.e., a preventive function. It is believed that a consultant service offered by clinical psychologists at strategic places in the community would reap infinitely greater benefits than the isolated treatment of individual children. A strategic place, although not the only one, is the public school which deals with children in their most formative stage on an almost daily basis.

School people in Ohio, aware of their responsibility in the developmental pattern of children, are asking for child study and guidance services in their schools. Depending on the community, they are ready and able to support such a service, but at the present time they are not able to obtain it because of the unavailability of competent personnel who recognize the preventive function of mental health and its group operation in the public school. The average clinician appears unwilling or unable to take a public health role.

This report offers a formulation of principles upon which a child study and guidance service might be provided to the public school. It offers a tentative training pattern within which adequate staff might be prepared. With some modifications, it is a pattern congenial to the basic training of a clinical psychologist. In our opinion, it is a vocational pattern which offers high promise to clinical psychologists with an interest in public health. It offers a most satisfying type of professional service in which the opportunities are relatively unexplored.

Basic Principles

I. The first principle is directed toward obtaining a moderate perspective toward the correct role of any special service in the school system irrespective of size of system, wealth of system, or the availability of special services. It is as follows: THE CENTRAL AXIS OF THE PUBLIC SCHOOL IS A LINE DRAWN FROM PARENT TO CHILD TO TEACHER TO ADMINISTRATOR. ANY SPECIALIST OR GROUP OF SPECIALISTS SHOULD DEVOTE HIS EFFORTS TO GIVING SUPPORT TO THIS BASIC AXIS IN ORDER THAT IT MAY FUNCTION MORE EFFICIENTLY. Thus, specialists and special services are all supporting the joint effort of administrator, teacher, child and parent. This is stated as a first principle to reinforce the concept that teaching is the function

*This paper is based on a report to be submitted to the Ohio Commission on Children and Youth by a committee of psychologists, psychiatrists, school administrators, directors of pupil personnel, guidance personnel and other related specialists under the chairmanship of the author.

of the school; that the specialist is a valuable accessory, not the central character on the stage of education. This principle does not depreciate the specialist; it merely defines his role and gives perspective to a total program.

II. This principle is based on the preceding one in that if specialists are to support the basic teacher-child axis they should exert that supporting pressure for all children. It says: CHILD STUDY AND GUIDANCE SERVICES SHOULD APPLY TO THE ENTIRE STUDENT BODY OF THE PUBLIC SCHOOL FROM PRE-SCHOOL TO POST GRADUATE. THEY SHOULD NOT BE LIMITED TO CERTAIN GRADES OR TO CERTAIN TYPES OF DEVIANT CHILD.

III. This principle is based on the idea that if special services are to support the parent-child-teacher-administrator axis, their advantages should not be restricted to a single pressure point on that axis (the teacher-child point). They should also help to strengthen other strategic linkage points involving both intra-personal and inter-personal problems. It says: THE SPECIAL SERVICE FUNCTION SHOULD BE BROADENED TO INCLUDE THE ADMINISTRATOR-TEACHER RELATIONSHIPS; TEACHER-PARENT RELATIONSHIPS; PARENT-CHILD RELATIONSHIPS; AND OTHER COMBINATIONS OF THESE BASIC POINTS OF INTER-PERSONAL AND INTRA-PERSONAL ADJUSTMENT.

IV. This principle is derived in part from the preceding one. It is based on the idea that he (the specialist) should be able to operate at more than one point and at more than one level. He should be concerned with and should influence certain general factors which if controlled would obviate the formation of many of the individual problems with which he has been accustomed to working on an individual basis. For example, the discussion of a single case with a teacher may provide sufficient insight into a common problem to prevent the occurrence of that problem in other children in her class. The specialist's concern with the planning of school-wide programs and services like curriculum building, seating, health, and others may be major contributions toward meeting the needs of many children simultaneously. Such broad efforts may obviate hundreds of hours of individual work of the traditional face to face type. The principle is stated thus: THE SPECIALIST'S ACTIVITY SHOULD NOT BE RESTRICTED TO THE SOLVING OF INDIVIDUAL PROBLEMS ON AN INDIVIDUAL BASIS IF HIS KNOWLEDGE AND SKILLS WOULD JUSTIFY MORE EFFICIENT APPLICATION AT A BROADER PREVENTIVE LEVEL.

V. SPECIALISTS SHOULD HAVE SPECIAL SKILLS AND KNOWLEDGE BASED ON APPROPRIATE AND COMPREHENSIVE TRAINING. This principle is a recognition of the outstanding single barrier to the obtaining of qualified specialists for service in the public schools. This single barrier and the complications which derive from it can often be pointed to as the reason why a given school system has a completely inadequate child study and guidance service and why IT MAY BE NEARLY IMPOSSIBLE EVER TO HAVE AN ADEQUATE SERVICE IN THAT COMMUNITY without rather drastic personnel reorganization. This barrier consists of the belief, rampant in educational circles, that the specialist should come from the ranks of the teaching or administrative group. In its worst form in child study and guidance circles, it has meant choosing an older teacher (sometimes too old for effectiveness in the classroom, the only area in which he has had training), and placing that teacher in an office with a fancy title of Guidance Counselor or even Director of Guidance. In a slightly less innocuous form, it has meant taking a teacher well liked by the children and setting aside part of his or her day to "counsel" with them. Thus we have "specialists" without specialties; a facade behind which lies a wealth of good intention and a minimum of competence.

This principle does not question the advantage of having a specialist in child study and guidance services who has had teaching experience. Such experience can provide an excellent backdrop for thorough training in the specialty. In any event, the specialist must have included in his training the opportunity to study and gain experience in the school setting in which he would practice his special skills.

VI. CHILD STUDY AND GUIDANCE SERVICES SHOULD BE UNDER UNIFIED DIRECTION AS OPPOSED TO A CONSTELLATION OF SEPARATE SERVICES WORKING WITHOUT CENTRAL CONTROL OR EVEN IN DIRECT OR PASSIVE CONFICT. This principle is based on the idea that all of the pupil personnel services can best give planned assistance to the central axis of parent-child-teacher-administrator if they are working in accordance with the general objectives of the school system. Conflict or ineffectiveness is less likely if the various services look for unified direction to a single person.

In the small system the child study and guidance specialist would very likely comprise the entire special service. Paradoxically enough, the small system needs the most highly trained man to carry out the total child study and guidance task in the absence of assisting colleagues who might specialize in given aspects of it. He would be responsible directly to the superintendent. In the slightly larger system where there is more than one specialist, one of them might act as the nominal head of the child study and guidance service as assigned by the superintendent. In the medium-sized system the superintendent may assign another administrator with good organizing and administrative ability to head up the corps of specialists. This administrative person might be designated as Director of Pupil Personnel. In an even larger system an administrator might be chosen with the rank of Assistant Superintendent in Charge of Pupil Personnel to organize and direct a unified Child Study and Guidance Service. The committee can give no definitive answer to the question of specialist versus administrator as the director of a program. Looking at the programs now in being, the practice in the larger systems is to draw the director from the administrative ranks. In fact, oftentimes the position of Director of Pupil Personnel is a stepping-stone in the administrative hierarchy.

Taking the long view in terms of programs not yet developed, it appears that the specialist would have to be the first member of a beginning child study and guidance service. He would, in fact, be a consultant to the superintendent. At this stage it would appear neither sensible nor efficient to put an intervening administrator between him and the superintendent. As he gathered about him other specialists to complement his own skills or to assume some of his activities, he might become the administrator of child study and guidance services; or the administrator might assign another person to administer the service and allow the specialist to doff the routine administrative duties that a growing department has forced upon him. Unfortunately, doffing administrative routines often means doffing some of your salary, since the administrative job conventionally carries the larger stipend. The eventual decision concerning who shall be "boss" is a function of the school system and of the personal capabilities of the men involved.

The Functions of a Child Study and Guidance Service and the Competencies Required to Carry Out These Functions

General Statement of Function. In basic principle number one, this committee expressed the central importance of the parent-child-teacher-administrator axis around which all special services must revolve of necessity. Implicit in this supportive role is the assumption that a special service brings to the central axis certain skills and competencies not present in the teacher administrative team to assist that team to function more efficiently.

One of the oldest, yet truest, slogans of education is that we take a child where he is and move him from that place forward. In its essence, then, the function of a special service is to offer methods and techniques to the teacher, parent and ad-

ministrator for the more precise determination of the current position of any child or group of children, socially, physically, and psychologically in order that appropriate steps may then be initiated to facilitate the forward motion of the individual or group along normal developmental pathways. In other words, the function of a child study and guidance service is to determine more precisely the differences that exist among children and groups in order that appropriate steps may be initiated to compensate for these differences and for the better arrangement of the total learning process.

Specific details concerning the function of a special service are offered here under two headings: I. Kinds of data and data-gathering techniques required in the study of children. II. Techniques in which competency must be gained if we are going to modify the behavior or conditions for behavior in children once data has been gathered and evaluated.

I. Kinds of data and data-gathering techniques required in the study of children should include:

A. A medical and dental examination involving medical personnel.

B. An evaluation of intellectual level through formal and observational techniques.

C. Formal, observational and historical achievement measures in terms of general and specific development.

D. Varying measures of personality involving projective, paper and pencil, and observational procedures.

E. Measures of "potential" using aptitude, interest, and preference techniques in the field of vocational "potential" as well as other specific behavioral areas.

F. Both historical and current evaluations of environmental home settings applying case study techniques and other sociometric techniques.

G. Evaluation of the school setting through observation, interview, and other measures.

H. Evaluation of the child's inter-personal dynamics; his relationship to self, to parents, to teachers, and to his peers. This involves his evaluation of his social setting as well as the evaluations of him by the people of his world.

I. A method for recording in appropriate usable form the data gathered above.

II. Techniques in which competency must be gained if we are going to modify the behavior or conditions for behavior in children once data has been gathered and evaluated.

A. Individual and group therapy from the most shallow "talking it over" to intensive personal therapy.

B. Therapeutic conferences with teachers, parents and others toward the eventual modification of the child's behavior.

C. Group discussions with teachers and administrators to aid in the solution of common problems.

D. "Manipulations" of various social agencies in the community to change actual physical conditions or social conditions of the children.

E. Manipulation of the curriculum to attack group problems of failure and adjustment.

F. Provision of special classes and assignment to special classes where appropriate.

G. The development of remedial programs for a child or group in various behavioral areas such as reading.

H. Vocational exploration for vocational choice.

I. Appropriate referral to consultant specialists.

Implications

The implications for the training of specialists for work in preventive mental health in the public school are moderately clear. In order to avoid fruitless argument among competing professional groups concerning the division of these functions this paper avoids purposely the naming of professional categories in advance. Competencies and requisite training are stated in terms of public school need. Practically, those responsible for graduate training are taking steps to incorporate the bulk of these skills in a single person with a prolonged training at master's or doctoral levels. It is obvious that the training of the clinical psychologist now encompasses many of these functions. Given a public health orientation, training in the interpretation of group data in a public school setting, an understanding of the school age as a behavior stage available for the efficient modification of individual and group behavior, the clinical psychologist is a "natural" to assume the central role. The writer predicts conservatively that within a few years the public school is going to provide a central platform from which clinical psychologists may make their most effective contribution to national mental health.

CASE STUDY OF AN ATTITUDINAL PATHOSIS

FREDERICK C. THORNE

University of Vermont

Introduction

Since the publication of the original paper (1) describing the *attitudinal pathoses* as a new clinical classification of behavior disorders, there have been requests for a detailed case study illustrating the basic etiologic mechanisms. After an intensive survey of record files, the following case has been chosen because it illustrates the basic etiologic mechanisms in an attitudinal disorder complicated by psychoneurotic reactions. Cases such as the following involve difficult diagnostic problems in order to differentiate primary and secondary etiologic factors in the disorder. The case reported here was deliberately chosen because it presents a problem in differential diagnosis concerning which it is possible to reach a number of different opinions according to the orientation from which the case is studied.

Counseling was carried on with this case over a period of sixteen weekly interviews. Verbatim recording was not available but detailed notes were taken relating to the most important details of the case history and of *core* and *central* attitudes expressed by the client. For the purposes of this report, the data have been rearranged in chronological order to provide a longitudinal overview of the development sequences involved. It is only after a careful consideration of the total life situation as it developed through the years, that an understanding may be obtained concerning the dynamics of the case.

Case Report

Identifying Data: The client (Mrs. D) is a 60 year old widow, born in New York State and who has lived most of her life in New England. She is of Anglo-Saxon ancestry and gives nominal allegiance to the Protestant Episcopal Church. Her education was begun in a public elementary school and completed with five years at a select girl's finishing school where she was considered a brilliant student. She married at age 18. She had three children. Her husband died about 15 years ago. Since then, she has supported herself as the proprietor of several businesses.

Behavioral Description: The client is a stylishly dressed elderly woman almost six feet tall. She has a heavy frame with strong hands, large feet and good muscular development. Constitutionally, she would be classified as a mesomorphic ectomorph. In spite of her rather masculine physical appearance, her personality and dress are decidedly feminine. She has a very interesting, intelligent face. Her features are well formed although large, and she has a very vivacious expressive manner. She is very friendly and hearty, smiling constantly, and gesturing with considerable histrionic ability. In her personal contacts, she is extraverted to the point of being almost overcordial. She always has some nice compliment to make, and expects to have it returned. The total impression is that we are dealing with a thorobred type, a person of high intelligence and culture, a person with abundant energy expressed in a vital interest in everything going on about her. She would be recognized as a member of the aristocracy in any social group.

Her conversation is brisk and stimulating. She tends to use superlatives in relating her experiences. Everything is "wonderful", "so beautiful", "such a marvelous mind", "I felt it *so* keenly", etc. Her manner carries something of the implication of royalty condescending to favor those of lesser standing. She is very cordial to everyone but there is nevertheless the definite awareness of her own constitutional and cultural superiority. This attitude of superiority is

not offensive because one has to admit that she is a "grand" person, and well able to live up to her own standards of superiority. One intuitively understands that here is a truly unusual person who would be at ease in the company of kings, and she has indeed actually entertained visiting nobility from Europe. One also gets the impression that here is a "driving" woman, a person who cannot help being a powerful influence in the lives of all who come near her because of her dynamic personality and the degree to which she implicates others in her activities.

One of the most significant traits of her behavior involves her manner of relating to other people. She makes a great effort to establish warm contacts with everyone about her. She approaches strangers with a cordial manner and seeks to involve them in the activities of the moment. In her role of host and "woman of the world", she is perhaps a little too aggressive in seeking to engage other persons in the activities which she considers proper. She tends to "play up" to people, and is hurt if they don't respond in kind to her. She shows a great need to "keep the ball rolling" at all times so as never to have a dull moment. At times, her reactions to people are a little too effusive in the sense that she tends to "gush over" in her compliments to them by saying "Oh, how wonderful you look!" or "Now sit right down and tell me what you have been doing. I want to hear every word of it." She has a "grand" manner of conducting her social relations, and everyone else is expected to fit in. For example, she will arrange cordial little "tête à têtes" in which she wants to know just how things have been going. Or there will be parties at which she is very jolly and gay and wants everybody to react in the same mood. She is quite insistent in getting people to do things her way. For example, she will come up and say "Don't you want to come over and join in the fun?" in a manner which leaves little choice. It is not easy to avoid playing her game. Her manner of approach is so personally direct and compelling that the other person literally has to go along with her if he wishes to avoid an open impasse. In this manner, she tends to dominate and organize social gatherings and almost to coerce others into her way of thinking. This trait has both good and bad aspects. On the one hand she is a really charming dynamic leader and a gracious hostess, and on the other hand there are many people who feel disinclined to go along with her style of things. Her children, for example, react violently against her "country-club" manner and her tendency to dominate social gatherings in a way which does not allow them to express their individuality. Her children somewhat resent her glamorous social personality which is displayed for the benefit of strangers. Even though her children are adults, they do not have much opportunity for conducting their social relations in their own way when she is around because she is so much more vigorous and forceful in automatically assuming the lead in every situation.

Her general approach to life is in terms of a "Super-Man" psychology. Her attitudes of superiority are probably not compensatory for deeper feelings of inferiority even though she may have such feelings from time to time. Her manner and behavior indicate a conscious idealism and perfectionism based on a deliberate philosophy of life in which achievement is the highest value. She perceives herself as a very intelligent, personable, gifted woman and her whole conscious effort is to realize her abilities to the utmost. She is constantly striving to be a *real* person. She makes great plans which she is usually able to accomplish. If she is unable to accomplish the highest things, it is only because she lacks the higher education necessary to provide her with the tools to work with. In general, she probably makes an excellent adjustment as long as she is associating with people of comparable levels of ability who can hold their own and establish tolerable dominance-submission ratios with her. Persons of inferior ability will probably feel themselves overwhelmed by the impact with her driving personality.

Psychometric examinations were not performed but clinical judgment rates her as of very superior intelligence. Intellectually, she shows an intensity and straightforwardness more characteristic of business men than of women. Although with only limited formal education, her wide cultural background permits her to hold her own with anyone. She thinks and acts with a directness unusual in women. Several times during the interviews she expressed the wish that somewhere she could find a man that she could look up to, i.e. who could dominate her. This will obviously be difficult because her own personality is so "driving" and domineering that it would require an extremely well integrated and superior person to keep her in her place. She has a large number of masculine friends but many of these appear to be submissive personalities with mother fixations. Many men do not like her because of her efforts to dominate every situation. However, this behavior is not too objectionable because of the attractiveness of her personality. If the other person stands up for himself, she is not overinsistent in having her own way. People who stand up to her get along very well after a working balance is achieved. She is very generous in helping people by giving her time or gifts. Whenever she goes visiting, she usually takes a large basket of dainties which she dispenses with a faintly condescending manner. Her son states that he never appreciated his mother's gifts because he felt that she used them to inflate her own Ego and to establish power over others.

Cause of Referral: The client was referred by a urological specialist who had been treating her for a minor bladder condition. Her chief complaint is persistent "pain in the eyes" of ten year duration which the urologist suspected of being psychosomatic in nature.

The client describes her condition as follows: "I am perfectly all right except for terrible pains in my eyes when I do heavy work. They came on about ten years ago. I don't know what started them. They just seemed to start all of a sudden. I have been to the best eye doctors in the country. I never got any help until I went to the Eye Institute where they found I have a rare condition of muscular unbalance. The doctors diagnosed it as due to spastic muscles causing painful accommodation. It bothers me so much that I can't use my eyes at all for hard work. I can't sit through a movie, or read a book more than 20 minutes without getting this pain and a headache. I can't read at all. I have a secretary to do all my bookkeeping. She also reads to me at night for an hour or two before I fall off to sleep. I don't know what I'd do without her . . . I don't know how to describe this pain in my eyes. The only thing I can say is that they just hurt. It's a steady dull pain in the whole eye. My eyes don't burn—they just pain. It's not a sharp stabbing pain. It usually lasts several hours until I can get some rest. Sometimes it lasts for several days . . . Nothing seems to help much. I have taken all kinds of medicines but it doesn't do any good. At the Eye Institute they gave me special exercises and glasses. They helped me some but not much . . . The only thing I can do is not to use my eyes for fine work such as reading or looking at the movies." Apart from this she is in very good health for a person her age, being able to work hard all day and into the evening running her business. She also reports minor complaints with her back, and cardiac palpitations on several occasions but she has not allowed these to interfere with her activities.

Family History: Both parents came from families who have been in America for several generations. Her father was a physician of international repute. The client says of him: "He was my ideal, a very attractive lovable person. Everybody adored him. He was one of the handsomest men I have ever known. When I was a little girl I used to be crazy about him. I would wait for him to come home at night to play with me. He used to love big parties but my mother didn't. I used to imagine that I was the hostess in our home. I liked to help my father

entertain at his parties. I just loved to sit and hear my father talk. He was the one who gave me my ideals. There wasn't anything he could not do. When he died, it was a great blow. I didn't know what I would do."

Speaking of her mother, the client says: "Mother was a Vassar graduate. She was very nervous and contrary as a child. As I remember her, she was intellectual, very beautiful but never cared about fixing herself up. She was quieter than Dad. She liked to go off by herself and write poetry. I don't think that she liked all the parties Dad had. He was a social lion but she never seemed to enjoy it. I guess Dad and Mother never got along too well in their later years. Mother always scolded. It seemed I always fought with her. Her voice was not very gentle. She was always irritated and in a fuss. It always seemed that I was taking Dad's side against Mother. I never seemed to be able to get very close to Mother."

The client had one older brother and one younger sister. She describes her sister as being "very lovely, all the things I was not." She did not have much in common with her sister. She compares her brother with her father. "He was also a howling success. He went to Princeton and did very well later."

The client states that she always had everything when she was a child. The family lived in a mansion in one of the best parts of the city. Her father's position placed the family on the highest social levels. They always had beautiful horses and carriages, and later, the best cars. The parents did everything possible to give their children everything. There was much pressure in the early home life for the children to be very successful.

Childhood History: "In childhood, I was a wild colt, hot-headed, not disciplined . . . Once I remember tying my little sister to a post box and putting a stamp on her to mail her away . . . I used to play hard at everything I did . . . When I was ten, I had typhoid fever with lymphatic swellings all over me. I hemorrhaged severely. After I got better my hair fell out and I got stout. I got to be very sensitive about my looks. I didn't want to be a little fat girl. I guess I overcompensated by trying to be the life of the party. My girlhood was not too pleasant. I was awkward and scrawny . . . I had a highly emotional girlhood. Sometimes I wanted to be somebody great. And then I would get very sad and want to be a nun or commit suicide."

"I didn't like it much at finishing school. I wanted to be home with Dad. I didn't want to go to Vassar either. I wanted to go to Europe to study so I wouldn't have to go to Vassar." Apart from these adolescent adjustment problems, the client's childhood appears to have been basically healthy and happy. She identified most closely with her father because of the similarity of their personalities, but the relationship appears to have been basically healthy.

Marriage: "I should never have married Chauncey Driscoll. He was a thorobred, intellectual, highly sensitive, exquisitely delicate in mind and body. I met him when I was 17 and I was married when I was 18. He was from a famous old New England family. His mother was a hypochondriacal invalid who always wanted all of her family around waiting on her. My husband had wanted to go to Cambridge but he couldn't get away from his mother. She was nice enough but the relationship was unhealthy. So he went to Harvard instead where he graduated *summa cum laude.* We became engaged when he was only 19. His family tried to break it up because he was so young. He loved me desperately. I've never been in love, I guess. I thought I was but I guess I wasn't. I loved him with pity to shield him from his own sensitivity. When I brought him home to meet the family, my father approved of him so I married him. I think my father wanted me to marry him. One day before the marriage, I almost had hysterics but nobody thought much of it. We had a grand wedding."

She became pregnant almost immediately after her marriage. "Chaunce always seemed uncertain of himself. He always wanted quiet solitude. He was

very sensitive. If someone else vomited, he would vomit too. I didn't want to have a baby so soon, but he wanted to. He said he wouldn't feel safe until he had me, until I had a baby by him. I guess he thought I might leave him and he wanted to pin me down. I didn't want to have a baby so soon. I was really sorry for myself. But I loved Betty when she came. I nursed her for only 3 months and then he got me pregnant again. How I resented that baby! I was just as sorry for myself as could be. I was always wanting to go to New York and have a good time. I remember that year I wanted so badly to go to a big New Year's Eve party in New York but Chaunce didn't want to go. I was furious. He settled that by getting me pregnant again."

Financial Security. Within three years after their marriage, her husband had attained fame and wealth from literary efforts. As the writer of best-selling novels, he became a popular social figure. "We led a life of glamor, high tension and adulation. We had too much money and social position." Within a few years, their income ran as high as $50,000 per year and they could have everything. However this was not a completely happy period for her because after the birth of her second child she hemorrhaged for more than a year and was largely confined to bed at home because of asthenia. Finally her father arrived on the scene and arranged medical examinations which resulted in the discovery and treatment of a pelvic condition. During this period, she still felt dissatisfied because her illness and home duties with the children kept her from joining with her husband in the activities into which his prominence gained him entrance.

Marital Rifts. During the period of his greatest literary and social successes, her husband began to drink with his friends. One night he stayed in town and contracted gonorrhea while intoxicated. On his return home, Mr. D was so obviously upset and conscience-stricken that Mrs. D suspected something and Mr. D confessed that he had contracted veneral disease. "At first I was heartbroken. Chaunce offered to let me divorce him but after thinking it over, I decided not to. After all, I had married him for better or worse and I decided I had to stick by him now. That was when the trouble really started. For some reason, it proved difficult to clear the condition up. He continued to have a discharge for more than six years before he finally got it cured . . . That time was surely Hell. We didn't have relations for more than six years if you can imagine two young people sleeping together for that length of time and not getting together. I tried to keep up a good front to outsiders. I used to flirt with some of the attractive men but I never went any further. Even that made me feel guilty. There was one man that I could have fallen in love with but he was older and nothing came of it. It was during these years that Chaunce began to drink harder and finally came to be an alcoholic. It got so I could never depend on him. One Thanksgiving we were going to have a lovely party but he turned up dead drunk and spoiled everything. I don't blame him now. I guess it was just the twenties when everything seemed to be wild."

It was during this period that Mrs. D became more active in running the family affairs. Almost on her own, she designed and had built a lovely country home. However this did not last long. One day she was sitting with the children in the living room when flame burst up before their eyes through the register. The house was gutted. Shortly after this experience, her oldest son Bob began to stutter. Mrs. D rebuilt the house in 1924. She complains that she had to assume all the responsibilities alone. "If only Chaunce had been able to help me some. But he had his writing. He was so sensitive and easily upset that it was easier to do it all myself. But we weren't really happy there. Chaunce was drinking more and more heavily. I decided that we ought to move away from New York and all the excitement and overstimulation. I heard of a beautiful estate that could be bought cheaply at Pinehurst, N. C. So we bought it and moved down. For a time things were better. After having been apart for six

years we finally got together and had Billie. But it didn't last. Little estrangements began to come between us. It was actually worse for Chaunce down there at Pinehurst what with all the parties and drinking. It seemed that there was a party going on all the time. Chaunce was drunk most of the time and he didn't get much writing done. I tried to be very gay. It was there that we met and entertained visiting royalty. But the life was too much for our health and our pocketbooks."

Europe. "Because of the whole situation I thought it would be wise to go to Europe and live. I had always wanted the children to be educated in Europe. I don't know why now, because now I know that American schools are better, but at the time I thought it was the only thing to do. At first we had a lovely time in France and Switzerland. We were able to lease a chateau at Cannes and I kept busy by starting a school for American children. For a while things seemed to be like old times again. Chaunce stopped drinking so much and was able to write again but he seemed to have lost his touch. Everything went well until about 1936 Mother came over to visit us. We were having a lovely time when one day we got a cablegram that Dad had died suddenly. As you would expect, Mother was pretty upset and it seemed that the only thing to do was for me to go home with her. I hurried around and got things organized, getting a nice young girl to be governess for the children. I don't know how we made it home on the boat. I don't seem to remember much about it but finally we got home. I stayed a few weeks and then returned to France. When I got back things were not going so well. Chaunce was drinking again and we seemed further apart than ever. I didn't know it then but I guess he had been living with the governess. At least she seemed to be closer to him than I was. By this time, things were beginning to get difficult financially. Chaunce was not getting much in royalties anymore. I had some money that Dad left me but that wouldn't last forever. Finally the money seemed to be running out so we had to come back home. We got home alright and tried to get settled down. Chaunce did not seem to be happy. I didn't know it at the time but he was still writing to the governess who was now in California. Apparently she was sending him money because he had promised to leave me and marry her. Probably Chaunce would have left me if he hadn't suddenly become so bad that I had to put him in a sanatarium. He seemed to be getting better when all of a sudden he contracted pneumonia. His last days were terrible for me. He got the children alone in his room and talked to them. I don't know what he said to them but I think he tried to turn them against me. They were terribly upset and would never tell me what he had said. Chaunce had apparently telegraphed the governess and she came on to make things worse. Well, to make a long story short, he died leaving me nothing except three children. It took me about three weeks to collect my wits and then I decided that I would have to go to work."

A Career. Showing great imagination and initiative, Mrs. D persuaded a large department store to open up a fancy specialty shop to cater to society women. This venture was a success almost immediately. Within two years she was earning $15,000 per year and able to establish a home again for the children. Betty, the oldest child assisted her as a personal secretary. Bob, the second child was in college, and Billie was still in high school. This was indeed a hectic period. In addition to her executive duties at the store, Mrs. D became involved in the social whirl. She organized a series of elite entertainments which were well attended. All in all, she did not have much time at home. The children complained that they never saw her. However, she was completely happy in achieving success on her own. If she could not give so much attention to the children, at least she was providing satisfactorily for their economic support and this was a great relief.

Tragedy Strikes Again. Mrs. D states that she always loved her children dearly and made every sacrifice to give them the best possible upbringing. They attended all the most fashionable schools and did all the proper things. Bob, the oldest boy, in thinking back over his childhood states that "we always did what she wanted us to do, never what we wanted ourselves." Following the death of his father, Bob got out of hand and Mrs. D could do little with him. He went away to college and never returned home again. He married in New York without his mother's knowledge, however she tried to patch the break by showering them with material gifts. Mrs. D. says: "I never went to see them without taking a big basket full of good things to eat." During this period, Betty married an attractive young athlete against her mother's advice. Mrs. D tried to dissuade her on the grounds that she could do better socially and financially if she waited a little longer. Betty eloped but made up with her mother immediately thereafter.

Mrs. D. says that she worried the most over Billie. "He was a sensitive child just like his father. The first thing I noticed was that he seemed to prefer to play alone at home by himself. I fixed up a lovely playroom for him with all the toys a boy would like but he never seemed to invite any friends home. He would just be waiting for me when I came home . . . I didn't think much of it at the time, but when he was a senior in high school he ran away from home. I didn't hear from him for several days. He apparently visited Bob in New York and got a little money and then started out for Florida. They picked him up there and of course I telegraphed money for him to come home. When he came home I said: "Billie, how could you do this to Mother?" He said that he didn't know why but told me that he loved me dearly and would never do anything like that again. The next year he went to a little college in the South. One day in the middle of the winter I came home from work and there was Betty waiting for me. She said: "Mother, something terrible has happened. Billie has shot himself." It was true. Billie had shot himself through the head. He hadn't given any warning. Shortly before he had sent me some lovely poetry he had written. We found some half-written letters which indicated that he had been desperately in love with a girl. He wasn't certain of her. I think he thought she was going to leave him. Anyway, something went out of me. I've never allowed myself to think of that scene again. It is too painful."

A New Life. Following Billie's death, Mrs. D entered a period of instability and confusion. She apparently began having difficulties with managing the specialty shop because she resigned her position. For a time she visited friends and then she secured temporary positions of several types. Gradually her grief lifted and she began to plan for the future again. She located and purchased a small resort business which also succeeded immediately because of her charm and ability as a hostess. She is happy and mentally occupied during the rush season when business requires her constant attention. During the off seasons, she has had more difficulty keeping herself occupied. Her most intimate contact is with her secretary-companion, a college girl who has been with her for three years. Mrs. D makes great demands on her secretary, who is kept busy during literally all of her waking hours. In addition to her business duties during the day, she acts as a companion to Mrs. D at night, reading to her or otherwise entertaining her. Recently, Mrs. D has had some little unpleasantnesses with her secretary who is beginning to act more independently. Mrs. D has criticized some of the secretary's friends, and otherwise has become overregulatory to the point where she stimulates rebellious attitudes. There has been considerable tension and it appears that the secretary may resign.

With her daughter, Mrs. D is maintaining cordial relations at least on the surface. Betty is apparently very loyal and loving with her mother. Betty now has two small children and a husband who is making a comfortable but not a generous living. Betty has been partially incapacitated by an undiagnosed

gastro-intestinal disorder which appears symptomatically to be mucous colitis. Betty is somewhat insecure and expresses feelings of inadequacy. Mrs. D attempts to help Betty wherever possible, but by mutual agreement she does not get too close. They get along very well at a distance.

Bob has turned out to have serious personality problems of his own. He has been undergoing psychoanalytic treatment for several years in order to "find" himself. Married to a vigorous young girl who is herself earning several thousand dollars a year in literary work, he does not feel capable of being a parent yet and therefore has had no children. He is holding a minor executive position in advertising work. Bob has a rather unhealthy personality. Although very intelligent and well educated, he has an unbearably stilted and pompous manner. He speaks with a painful stiffness and affectation, using an accent which is an ill-chosen mixture of Harvard Boston and British styles. Ideologically, he appears very much confused and over-inhibited. He obviously feels rather inadequate in relation to his wife who is a very vivacious person. He attempts to control her by an insufferable manner of authority and dictatorialness. Bob's conflicts in relation to his mother are very close to the surface. His attitudes are revealed in a little incident related by his mother which occurred within the last year. Mrs. D had driven down to the city to see Bob and his wife. On her arrival she was attempting to park her car in a rather limited space when she inadvertently rubbed fenders with the car ahead. Somewhat impatiently Mrs. D. said: "Oh, Bob, why can't you help me?" Whereupon Bob became very angry and retorted: "Damn you, you old bitch, you may have your own way, but you can't have me too." This episode upset Mrs. D deeply. She finished her visit politely but resolved that she would not return again until Bob made some apology or changed his attitude. This he has shown no signs of doing.

During the last year, Mrs. D. has been living quietly carrying on her business. She does not mix socially much in town. When she first moved to this locality, she was very active in trying to meet people socially and to influence town business. However, the natives reacted to her efforts with suspicion and did not cooperate with her activities. Shortly after her arrival, Mrs. D had as one of her guests a French countess who she tried to introduce about town. Apparently the townspeople misconstrued her efforts, thinking that she was attempting to rise above them socially. Now, Mrs. D is polite to her neighbors but does not have much to do with them.

Course of Therapeutic Interviews

Roughly the first half of the counseling, or eight interviews, was taken up with the eliciting of Mrs. D's past history. At the beginning, her attitude was one of condescending superiority. She made frequent references to the professional standing of her father, and to the many prominent eye specialists that she had consulted. She let it be known that ordinarily she would have consulted a prominent psychiatrist in Boston but had agreed to come to us locally only because of the very good recommendation made by her urological consultant. During the early interviews, she made a definite effort to dominate the situation by relating the details of her social background and previous successes. During these early interviews her productions consisted mostly of factual statements concerning the details of what had happened to her. She continued to paint a very rosy, self-contented picture of her life and accomplishments. Although most of this material was reacted to nondirectively, the author felt intuitively that Mrs. D needed to feel that the therapist was a superior person, on her own social level, and more than her equal. Accordingly, we acquainted her with enough of our own professional background so that she was able to feel confident that she was indeed consulting a specialist. The therapist made a number of clever interpretations worded to appeal to her intellect. Gradually, she dropped her social affectations and began to reveal her inner feelings.

The last seven interviews were given over to the analysis and interpretation of the major events of her life. With the exception of the relation of her eye condition to the suicide of her son Billie, most of her emotional conflicts were close to the surface and it required little interpretation to give her insight. A major portion of the therapeutic effort was directed toward ideological reorientation relating to values in life. Her extreme idealism and perfectionism were discussed both in relation to herself and what she could expect from others. She easily understood the fact that her own "driving" personality could be very threatening to those of lesser resources. Starting with her relationships with her secretary, the details of her interpersonal contacts were analysed in detail. Even within the four month period of therapy, Mrs. D was able to improve her social contacts markedly. After rapport had been established, she was very amenable to suggestions and made a successful intellectual attempt to handle her affairs on a more quiet and relaxed level.

Her "pain in the eyes" showed no improvement until she was able to associate more freely concerning her relations with Billie and his suicide. For ten years she had been rigorously repressing any conscious thoughts of his death. She now stated that she had thought of the suicidal scene as being too horrible to visualize. "I could not bring myself to think about it. I could not bear to imagine the bloody scene." After pointing out the temporal connection between this tragedy and the onset of her symptoms, the therapist interpreted the pain in the eyes as a psychosomatic symptom possibly symbolizing her unconscious rejection of the painful memories related to Billie's suicide. Following this interpretation, and a general improvement in her overall adjustment and with particular reference to allowing more freedom to her secretary, the eye symptoms began to improve until Mrs. D began to be able to read or attend the movies without trouble.

The therapeutic effort was terminated after an interview with her son Bob, in which a situational adjustment was made which would permit the reestablishment of at least superficially cordial relations between mother and son. After this had been accomplished, Mrs. D said that she did not feel any need to come for further interviews.

Mental Status at End of Therapy. At least in her contacts with the therapist, Mrs. D was markedly changed after four months of counseling. The most apparent alteration has occurred in the attitudes which characterize her way of life. She has been led to reevaluate some of the basic attitudes upon which she had guided her conduct intellectually. First, she is more aware of the feelings of others, and of the importance of not attempting to regulate everyone in her environment. Second, she is less perfectionistic and more realistic in her expectations concerning herself and others. Third, she is less concerned about matters of social standing. Finally, her entire system of values is more appropriate to the conditions under which she must live. This has come about by the analysis of many little situations in which her adjustment had not been good.

Emotionally, she is chastened by the understanding of her complicity in the tragedy of her family life. However, she is reacting to this intelligently. The interpretation was offered that any errors that were made were honest mistakes for which no one was to blame since everyone acted as they did with the best motives and that the eventual outcomes were the results of human frailty.

Diagnostic Formulation

This case presents an interesting problem in differential diagnosis. It is immediately apparent that the presenting complaint, pain in the eyes, represents only a secondary psychoneurotic reaction which is relatively insignificant against the broader background of her entire developmental history with particular reference to her whole family life. There is such a complexity of pathogenic factors at all periods of the life history that it is difficult to assign proper weights in attempting to evaluate the whole situation. Although the case includes much material suscept-

ible of psychoanalytic interpretations, the Law of Parsimony will be observed drawing only such conclusions as seem valid in terms of objective evidence. In order to review the major facts of longitudinal development, it is desirable to outline the salient features of her life history.

Mrs. D was born with good heredity into a well-to-do upper class family which offered every form of material security to the children. The family was dominated by a brilliant handsome father, an internationally known physician, with whom Mrs. D identified strongly throughout her childhood. Unquestionably Mrs. D derived her attitudes of perfectionism and idealism from her father. Mrs. D apparently competed with her mother and sister for the father's affections, and there was probably a good deal of hostility between Mrs. D and her mother and sister. As a child, Mrs. D was very active, vigorous, vitally interested in everything, alert, intelligent, sensitive, emotional, headstrong and already beginning to dominate everyone about her because of her natural ability, drive and forcefulness of personality. Apart from some self-consciousness over losing her hair and obesity after an illness at age 10, her childhood appears to have been within mormal limits.

She married at 18 after having been rushed off her feet by the impetuosity of her husband and without having developed any real love for him. He was a brilliant, sensitive but insecure person who quickly made a small fortune as a writer. At that time, Mrs. D was still very immature and wanted to have some fun and excitement before settling down. Instead, she became pregnant twice in rapid succession and shortly thereafter her husband contracted a venereal disease which caused them to discontinue relations for six years. During the early years of their marriage, Mrs. D's dominant personality apparently constituted a serious threat to her husband who began to drink heavily. During the next 15 years, Mrs. D gradually assumed more and more responsibility for making all decisions of the family. This was the period of the "roaring twenties" when it seemed as if all the values of life were warped. Looking back Mrs. D later saw that she had been overinfluenced by the false values of social position, trying to overregulate her children to make them into something which they were not, excessive drinking and party life, and increasing incompatibility with her husband. Adjustment problems became more difficult following the stock market crash and depression, the death of her father, the involvement of her husband with the children's governess, and finally his premature death from alcoholism leaving her almost destitute. In reviewing this period, Mrs. D has insight into the fact that with his weaker, unstable personality, her husband had been unable to feel secure with her. Her driving perfectionistic attitudes had apparently exerted an intolerable pressure from which he escaped into alcoholism.

On becoming widowed, Mrs. D entered a difficult period of life in which she not only took care of the three children but single-handedly recouped her fortune in a brilliant financial success. Here she demonstrated her basic ability and vigor by embarking on a succession of business ventures in which she was successful. However the accumulated tensions of an unstable overregulated early life began to reveal themselves in the form of emotional instability and neuroticism in her children. Her youngest boy first ran away from home and later committed suicide. Mrs. D apparently had much unconscious guilt because shortly thereafter she developed the conversion symptom of "pain in the eyes" because she could not bear to visualize the scene of her son's death or even think about him. Her relationships with her daughter became temporarily tense when the daughter married an unacceptable person but later made penance to her mother by increased loyalty and overt affection. Her relationship with her oldest son became progressively more strained until he broke away from her in an uninhibited burst of aggression and hostility whose signifience was so patent that Mrs. D could not overlook his rejection of her.

In summary, Mrs. D has always been a very strong, constitutionally superior, emotionally intense, intellectually gifted person. Unfortunately she did not have an education commensurate with her natural abilities. Her style of life was based on a "Super-Man" philosophy derived from her idealistic attitudes toward her father. She is what might be called a "driving" woman; her personality is so vigorous and strong as to place intolerable pressures on those who are unable to keep up with her. It is inescapable that her personality has exerted important effects upon the lives of her family. Her husband died prematurely of alcoholism. The oldest son was in rebeilion with his mother for many years but was unable to break away from the resulting conflicts even after psychoanalysis lasting three years in adult life. He is still in a state of acute conflict manifested by open hostility which his mother cannot ignore. The daughter maintains loyal relations with her mother but has a psychosomatic condition which physicians have been unable to diagnose or treat. The youngest son withdrew progesssively, and developed such conflict and inner tension that he ran away in high school and later committed suicide. Mrs. D, herself, has survived these traumatic experiences remarkably well with only the development of a relatively mild conversion symptom and a mild sense of frustration and unhappiness.

Integrating all these facts, it may be concluded that Mrs. D herself is a relatively healthy personality inspite of the pathogenic effect of her attitudes upon others. At no time in a very strenuous and difficult life has Mrs. D been incapacitated psychologically except by a minor, though subjectively important, conversion symptom of "pain in the eyes", and even this symptom did not appear until after the trauma of the suicide of her youngest son. Although her conversion symptom is psycho-

neurotic in nature, it appears to be a secondary personality reaction. The primary etiologic factor appears to be in Mrs. D's ideological orientation toward life as revealed in her attitudes. It is interesting to speculate what she might have accomplished with an education commensurate with her native ability. As it was, she married at a very early age when many of her attitudes were very immature. It should not be assumed that Mrs. D's attitudes are the result of blind unconscious emotional impulses. On the contrary, Mrs. D has a very controlled, deliberate, rational intellect. Her planning has been deliberately based on what she thought was best. Her early life, and particularly her associations with her father, conditioned her in attitudes placing high value on (1) being a very superior person, and (2) striving for social status including having the "right" kind of friends, going to the "proper" schools, travelling in Europe, and "doing" all the right things. If any one *core* attitude can be distinguished it is concerned with being somebody, amounting to something, being a real person in all senses of the word. There is nothing abnormal or unhealthy in this style of life but only in the patterns developed for achieving it. With a better education, particularly in the matter of influencing people close to her, she might have turned out to be a very adequate wife and mother. As it was, she was a member of a generation which placed great emphasis on regulating people to make things come out right. Mrs. D felt that she *had* to do what she did because it was the only way she knew to "bring up" her children properly. The most unhealthy thing about Mrs. D is her ideological orientation to life which is expressed in a great many attitudes each reflecting some untenable value concerning how life should be lived. The superficiality of many of these intellectual attitudes is revealed by the ease with which she was able to change her way of handling people after she had gained insight into what she was doing wrong.

The final diagnosis recorded in this case was "Attitudinal pathosis, of mixed type, with secondary psychoneurotic conversion symptoms, and severe situational maladjustments of family life." It is hypothesized that the primary etiologic factors are the attitudes which reflect her ideological orientation to life.

Discussion

In differentiating attitudinal disorders from psychoneurotic disorders, an important criterion is whether the unhealthy attitudes have been acquired and acted upon deliberately, rationally and through conscious volition, or whether they are determined by unconscious, involuntary factors. Attitudes may be acquired in a variety of ways, some within normal limits and others definitely pathological. *Normal* attitudes are derived through the ordinary conditioning of life, are held on conscious levels on rational grounds, and are capable of being abandoned voluntarily if the person rationally discovers their untenability. *Psychoneurotic* attitudes involve involuntary deep affective or physiological reactions, held on unconscious levels on irrational affective-impulsive grounds, and are very difficult to control rationally and voluntarily. It should not be assumed that all attitudes are determined emotionally or unconsciously. On the contrary, *normal* attitudes, opinions or beliefs are learned and may be unlearned in terms of their tenability or validity. If a person discovers that a particular attitude or belief is invalid, he abandons it on rational grounds. Under normal conditions, every person's attitudes are undergoing a constant process of modification and revision in the direction of greater internal consistency with assimilation of more tenable attitudes and elimination of invalid ones. The criterion in differential diagnosis is to discover whether an attitude is determined by normal conditioning or by affective-impulsive psychoneurotic factors in personality.

In the case here reported, it is postulated that we are dealing with a basically healthy personality in which mild psychoneurotic symptoms did not develop (except as a personality reaction) until after a final gross trauma was added to many years of cumulative stress. In terms of the criteria of social competency and lack of incapacitation, this client has always been of more than average social competency even

during the period when stresses and psychosomatic complaints were most intense. The background for the development of this life pattern was, of course, determined by the family situation into which the client was born, and also by the client's constitution and temperament. In terms of the early environmental influences, and the personality type with which we are dealing, it may be concluded that the pattern of development is also basically healthy with the exception of general attitudes and style of life which the client early acquired. It is interesting to speculate how her life might have been different if she could have had an education comparable to her level of intelligence. As it was, she did the best she could with the resources available. Considering that the *map* upon which she based her course through life was invalid in many of its attitudes and beliefs, it is remarkable that she did as well as she did. Her failures may be largely explained on the basis of the invalidity of the beliefs and values upon which she made rational decisions. Although affective-impulsive factors unquestionably exerted their influence, psychoneurotic factors do not appear to have played a very important role except in the form of secondary personality reactions to great stress and traumatic events. Evidence for the conclusion that the primary etiologic factors consisted of attitudinal or ideological inconsistencies operating on intellectual-rational levels of personality is adduced by the closeness to consciousness of her conflicts and by the relative ease with which a reorientation was accomplished by rational approaches. This client does not fit the picture of the typical psychoneurotic governed by uncontrollable affective-impulsive factors and showing the typical symptoms of neurotic incapacitation. She shows mild conversion symptoms, but study of the case history reveals that these are secondary and reactive to the primary attitudinal disorder.

Summary

A detailed presentation has been made of a case diagnosed as "Attitudinal Pathosis, of mixed type, with secondary psychoneurotic conversion symptoms and severe situational maladjustments of family life." The case is discussed from the viewpoint of differential diagnosis in order to properly weigh etiologic factors in the disorder.

Reference

1. Thorne, F. C. The attitudinal pathoses. *J. clin. Psychol*, 1949, 5, 1-19.

PARENT-CHILD SIMILARITIES IN PERSONALITY DISTURBANCES

E. LAKIN PHILLIPS

*Arlington (Va.) Guidance Center**

INTRODUCTION

The purpose of this paper is to bring additional evidence to bear on one basic tenet in clinical work with parents and children, viz., that children's psychological difficulties grow out of and reflect the parent's psychological difficulties. The general possibility of such a relationship existing between parents and children is most reasonable when we consider the parent-child relationships in such variables as intelligence (6, 8), trait measurements (8, 14), the gross number of mother's fears and the gross number of children's fears (5), plus the fact that children are in many ways the products of their parent's biological and cultural backgrounds (2, 3, 9).

Representative clinical studies (1, 4, 10, 12) attest, also, to the parent-child relationships in personality characteristics where problem children and their parents are considered. Among the findings in these research studies, one finds: Mothers with unhappy childhoods have children with a variety of neurotic traits (4); children who were nagged, criticized, punished and supervised strictly by their parents were judged by their classmates as being quarrelsome, disobedient, quick-tempered and nervous (1); the child of dominant parents tends to be nonaggressive and insecure, while children of submissive parents tend to be rebellious and aggressive (12). As early as 1909, Jung observed the subtleness of parent-child relationships and wrote, regarding the developing child, "What most influences him is the peculiarly affective state which is totally unknown to his parents and educators. The concealed discord between parents, the secret worry, the repressed hidden wishes, all these produce in the individual a certain affective state which slowly but surely, though unconsciously, works its way into the child's mind, producing therein the same conditions and hence the same reactions to external stimuli . . . " (7, p. 246). Sperling found among twenty children suffering from psychosomatic illnesses that, "Mother and child represented a psychological unit in which the child reacted to the unconscious need of the mother with correspondingly unconscious obedience; it was as though the child were given a command to get sick, which meant in reality, to stay dependent and helpless" (10, p. 377). Thompson (13) remarks broadly on interpersonal relations theory and, along with Sullivan (11), offers some theoretical discussions which cover the parent-child configuration as well as other two-person relationships.

METHOD

The files of the Arlington (Va.) Guidance Center were examined over a period covering the past two years and fifty-two cases, from among over one-hundred cases, were found which included relatively complete psychological studies of the children and records of extensive psychiatric interviews with the parents of these children. These data included a battery of psychological tests, play-sessions, and interviews with the children; and a series of interviews by social workers and psychiatrists with the parents. Of the 52 cases, 25 were boys between ages 5 and 13, and 27 were girls of the same age range. All children were at least normal in intelligence, and none of the children came from families earning less than $3,000.00 per year.

RESULTS

These data were treated in two ways. First, in terms of the presenting complaints (symptoms) culled from all the data on parents and children. In all, there

*Conducted by State Department of Mental Hygiene and Hospitals.

were 42 symptoms which occurred 132 times among the children and 81 and 52 times, respectively, among the mothers and fathers of the children. Seventy-one percent of the symptoms showed overlap (communality) among mother-child pairs, and 48 percent of them showed father-child communality.

The second way in which these data were treated was to compare the parents and children on the basis of the main complaints which the parent(s) saw in the child. On this basis, 40 out of the 52 families (77 percent) showed striking similarities. This finding covers an important qualitative result which cannot be quantified on the basis of the present data: The complaints regarding the child which bothered the parents most (their own evaluations), and which probably constituted the real basis for referring the child, were centered around the unresolved psychological difficulties which the parents had from their own histories. Three brief case histories will illustrate this point.

"Tom" age six, was referred because of his enuresis, and this was the behavior which concerned the mother most. The interviews finally brought out that the mother had been enuretic until she began to menstruate, that " . . . people used to kid me about it." The mother recalled in this context that as a child she had been so afraid that she would wet the bed that she refused to go to bed at night, or when she did go to bed she would lie awake for long periods of time disturbed over possible bed-wetting.

The mother referred "Curley," a girl of seven, around problems of discipline and control. In play-sessions, Curley showed much concern for orderliness, keeping her skirt spread out neatly when she sat down " . . . so it would not bunch up underneath me." In her doll-play, she said, "I have just ten minutes to cook dinner . . . just five minutes left!" And ". . . they (the dolls) can play for only one hour . . . I must tell Susan (the doll) when its time to do everything," and " . . . there, that's the way I want her to sleep!" The mother was extremely compulsive in her household chores, chastising herself in the interview for not keeping the house clean at all times, for not having meals ready on time, etc. Trouble with the child, then, was centered around the mother's demands on the child which were highly inappropriate for a seven year old, but which also reflected the mother's problem at the same time. The compulsive and irrational controls she placed on herself were communicated to the child and disrupted the child's behavior, thus making her a "problem" child in the mother's eyes.

"Karen," age eight, was aggressive, hostile and defiant, in her mother's and the school's reports. Karen repeatedly stated that she was "mad as the dickens," and was going to "tell off" everyone. In response to the S-B item, Opposite Analogies (L, VII, 5), "The point of a cane is blunt; the point of a knife is ——," Karen said, "Sharp enough to cut somebody's head off with!" with great vehemence. The interview series between the social worker and the mother showed how the mother dealt with threats in the same way, thus, "If I don't like people, I don't care how I treat them," and "I just spit out whatever comes in my mind."

Discussion and Conclusions

These data and the brief case materials illustrate the hypothesis on which this paper is based: Parent-child psychological difficulties arise out of a context of unresolved difficulties from the parent's behavioral history. The child's problems are apparently 'created,' not out of a context which defers to the question, "What's wrong with my child?" but out of a context in which the parent's disturbances in various areas of living are communicated to the child and get the child involved in problem behavior in his own right. Stated another way, we may say that children have to use their parents as models for adjusting to the multitudinous demands of living; if the parents are disturbed in some of their living relationships, this increases the burden the child carries and increases the likelihood that he will be involved in problem behavior in the same areas and in the same ways as his parents.

References

1. Anderson, J. P. *A Study of the Relationship Between Certain Aspects of Parental Behavior and Attitudes and the Behavior of Junior High School Students.* New York: Bureau of Publications, Teachers College, Columbia University, 1940.
2. Davis, Allison. Child Training and Social Class. In Barker, Kounin, and Wright (Editors), *Child Behavior and Development.* New York: McGraw-Hill, 1943.
3. Davis, Allison and Havighurst, Robert S. *Father of the Man.* Boston: Houghton-Mifflin, Co., 1947.
4. Field, M. Maternal Attitudes Found in Twenty-Five Cases of Children with Primary Behavior Disorders. *Amer. J. Orthopsychiat.*, 1940, 10, 293-310.
5. Hagman, R. R. A Study of Fears of Children of Pre-School Age. *J. Exp. Educ.*, 1932, 1, 110-130.
6. Hurlock, Elizabeth B. *Child Development.* New York: McGraw-Hill, 1950.
7. Jung, Carl. *The Association Method. Amer. J. Psychol.*, 1910, 21, 219-269.
8. Olson, Willard C. *Child Development.* Boston: Heath & Co. 1949.
9. Phillips, E. L. Intellectual and Personality Characteristics Associated with Social Class Attitudes Among Junior High School Students. (In press, *J. Genet. Psychol.*, 1950).
10. Sperling Melitta. The Role of the Mother in Psychosomatic Disorders in Children. *Psychosomatic Medicine*, 1949, 11, 377-385.
11. Sullivan, Harry S. The Meaning of Anxiety in Psychiatry and in Life. *Psychiatry*, 1948, 11, 1-13.
12. Symonds, Percival. *The Psychology of Parent-Child Relations.* New York: D. Appleton-Century, 1939.
13. Thompson, Clara. *Psychoanalysis: Evolution and Development.* New York: Hermitage House, 1950.
14. Thorpe, L. P. *Child Psychology and Development.* New York: Ronald Press, 1946.

THE CONSTANCY OF THE ALTITUDE: A NOTE

LEONARD DILLER AND ROBERT M. BEECHLEY

Northern New Jersey Mental Hygiene Clinics

Introduction

Recent studies[2, 3] offer evidence that the *altitude*, defined as the greatest success an individual achieves in a series of psychometric scales, might offer (1) a more valid index of an individual's "intelligence" than the IQ, and (2) a more adequate reference point for scatter studies than does the IQ. If the altitude is to be used in daily clinical practice, one of the minimum requirements it must fulfill is that it be at least as constant as the IQ. "The thesis that altitude is a valid measure of native capacity and not merely another conventional operational statistic depends in part upon the demonstration of its intra-individual constancy"[6]. It is the aim of this study, therefore, to compare the constancy of the altitude with the constancy of the IQ. A secondary aim is to demonstrate the application of the altitude concept to the Stanford-Binet Scale. Until the present, studies in altitude have been confined to the Wechsler-Bellevue Scale[6], and the Psychometric Patterns[4], both of which differ from the Stanford-Binet test in that they are composed of subtests made up of homogeneous point scales, whereas the Stanford-Binet is not.

Method

One hundred subjects with a mean age of 7.4 years, ranging from 2.6 to 13.25 years, were tested and retested on the Stanford-Binet (L) 1937. The subjects were patients who were referred to the Northern New Jersey Mental Hygiene Clinic because of school retardation, behavior problems, mental deficiency or determination of eligibility for adoption. The mean interval between test and retest was 2.1 years with range of .5 years to 6.25 years. Two indices of altitude as well as the IQ were used. In one method, the altitude was computed by translating the highest single success into an altitude age, e.g., a subject whose maximum was achieved in passing a single subtest at the eight year level received an AA (altitude age) of 7.2, although his MA as traditionally computed as the sum of successes may have been 6.0. The altitude was also computed by translating the three most successful performances into AA's and averaging them. In the second method a person who achieved a single success at the eight year level (AA = 7.2) and two successes at the seven year level (AA = 6.2, 6.4) obtained an AA of 6.7. The altitude age was changed to an altitude quotient (AQ) by dividing by the CA (chronological age).

Three measures of comparison of the constancy of the IQ and the AQ's were utilized: (1) the means and standard deviations of the IQ were compared with the AQ's on test and retest; (2) the differences between test and retest for AQ's and IQ were compared; and (3) the IQ's for test and retest were correlated, and the AQ's for test and retest were correlated.

Results

Table 1 indicates that none of the methods employed demonstrated a significant difference between the quotients on the retest and the quotients on the original test. All three indices appear to be fairly stable measures. The averages of the quotients are within 2 points of each other. The altitude computed by means of the first method averages 6-8 points higher than does the IQ, while the altitude computed by means of the second method averages 3-4 points higher than the IQ.

The average differences between test and retest are approximately 7 points for the IQ and 8-10 points for the AQ's. These differences are somewhat larger than would be expected in a non-clinic population[5], but do not differ significantly from each other.

TABLE 1. COMPARISON OF IQ AND AQ'S ON TEST AND RETEST

	Test		Retest		Difference		
	Mean	S. D.	Mean	S.D.	Between Means	CR	P
IQ	79.45	14.7	81.40	17.5	1.95	.08	<.90
AQ^1	88.0	17.9	88.5	18.45	.5	.02	<.90
AQ^3	83.6	17.03	84.6	18.22	1.0	.04	<.90

TABLE 2. A COMPARISON OF THE DIFFERENCES BETWEEN TEST AND RETEST QUOTIENTS

	Means of Differences	S.D.	Differences Between Quotients	CR	P
IQ	7.18	5.97	2.8 (IQ – AQ^1)	.32	.75
AQ^1	9.98	6.3	1.4 (AQ^1 – AQ^3)	.17	.87
AQ^3	8.5	6.23	1.32 (IQ – AQ^3)	.15	.88

Using the Pearson product-moment method, it was discovered that the IQ's correlate with each other on test and retest r = .820, σr. 033. The AQ's yield according to the single highest success method an r of .775, σr .4 and according to the three highest successes method an r of .785, σr .038. Tests for significance[1] of the differences between r's reveal that the CR of the difference between r's of IQ and AQ^1 = .84 and the CR of the difference between r's of IQ and AQ^3 = .78. Both differences are not significant and the r's may be considered as arising from the same population.

DISCUSSION

In applying the concept of altitude to a point scale such as the Wechsler-Bellevue, Jastak has utilized the average of the three highest subscales[4]. However, the authors found that for the Stanford-Binet scale, cases arise with only slight scatter in the test records so that this index comes out exactly the same or even below the IQ. For example, a child whose basal age is 7.0 and who has passed only one test at the eight year level would receive an MA of 7.2, whereas the AA if taken as the average of the three highest scores would equal 7.0. For this reason the highest single test on the Stanford-Binet was chosen as an alternative altitude measure.

Since the differences between the means of the IQ group and the means of the AQ groups are not significant (difference between IQ and AQ = 8.55, CR = .37), it may be argued that the reason why our results show no difference in IQ-AQ constancy is because of a lack of scatter in the test performances, i.e., the reason why the AQ holds up as well as the IQ is because both are measuring the same phenomenon. To obviate this objection, it was decided to divide the cases into two groups, so that in one group the AQ was to be separated from the IQ by more than 8 points, while in the other group the separation was to be less than 8 points. With 50 in each group, and a mean difference of 14 points between AQ^1 and IQ for the first group and a mean difference of 3 points for the second, it was discovered that the AQ still proved to be as constant as the IQ in correlations of test and retest (r's ranged from .82 to .88 with no significant differences between them). Our results would appear to indicate that the use of an AQ cannot be ruled out on the grounds that it is less reliable than the IQ. Although there may be exceptional cases where the AQ is influenced more by accidental extraneous factors than the IQ, it appears that for practical purposes the altitude is as constant as the IQ.

Some implications of our results are that (1) if the altitude quotient is to be used as the criterion for feeblemindedness in place of the IQ we should use a quotient several points higher (3 to 8) than the traditionally accepted quotients; (2) it would be possible to utilize the Altitude concept with performance tests and combine the results with the Stanford-Binet in order to get an altitude measure based on both tests. Two problems posed by the relationships between AQ and IQ are left untapped by our study: (1) It is possible to investigate the relationships between IQ and AQ before and after therapy. According to Jastak, the IQ should tend to rise in relation to the AQ which remains constant, because it is more apt to be influenced by personality disorganization. (2) It is possible to establish AQ - IQ ratios for the intellectually less endowed so that true feeblemindedness may be distinguished from pseudo-feeblemindedness.

Summary

A study of one hundred clinic children who were tested and retested on the Stanford-Binet demonstrates that the altitude age derived from the single highest success or the three highest successes, when converted into a quotient by dividing by the CA, appears to yield as constant and as stable a measure as the IQ. This stability appears to be independent of the size of the difference between the AQ and the IQ.

References

1. Guilford, J. P. *Fundamental Statistics in Psychology and Education.* New York. McGraw-Hill, 1950.
2. Jastak, J. Problems of Psychometric Scatter Analysis. *Psychol. Bull.*, 1949, 46, 177-97.
3. Jastak, J. A Rigorous Criterion for Feeblemindedness, *J. Ab. Soc. Psychol.*, 1949, 44, 367-378.
4. Jastak, J. Psychometric Personality Traits. *Delaware State Medical Journal*, 1949, 21, 165-176.
5. Terman, L. and Merrill, M. *Measuring Intelligence.* New York: Houghton-Mifflin Co., 1937.
6. Whiteman, M. Altitude as a Reference Point in Scatter Analysis. *J. Clin. Psychol.*, 1950, 6, 160-164.

A NOTE ON THE LIMITATIONS OF THE USE OF THE BENDER GESTALT TEST AS A DIAGNOSTIC AID IN PATIENTS WITH A FUNCTIONAL COMPLAINT*

LEO J. HANVIK

Colorado A. & M. College
Fort Collins, Colorado

It has been observed that the Bender Gestalt Test, although valuable as an aid in the diagnosis of the psychoses, and useful in the detection of brain damage, is virtually useless as an aid in differentiating between the "normal" individual and the individual with one of the more mild personality disorders (i.e. the neurotic). In the course of an intensive psychological study of a number of male in-patients admitted to a Veterans Administration Hospital Neuro-psychiatric Service and found to be suffering with functional backache, there was an opportunity for statistical comparison of the Bender protocols of a neurotic group (the patients with the functional complaint) with the protocols of a "normal" group (patients who were likewise admitted to the hospital with backache but in whom the diagnosis of herniated intervertebral disc, surgically verified, was ultimately made). The nature of the comparisons and the results thereof are reported briefly herein.

The Bender Gestalt Test was administered in the usual manner. It was given to a group of 23 patients with "functional" back pain and to a control group of 14 patients with a proven organic disease of the back. It should be noted that the patients in these two groups were markedly different clinically, the functional cases showing many neurotic signs, the other cases few or none. There were also outstanding psychometric differences (e.g. on the MMPI) which were highly significant statistically and which suggested that features of the conversion hysteria type were the salient characteristics of the functional group.

In studying the protocols three methods were used: (1) a "tally count" method, essentially an attempt at scoring the Bender, where a list was drawn up of 19 characteristics believed to show some variability from subject to subject. [1] The protocols were all examined by two clinical psychologists and rated for the presence or absence of each characteristic. A *t* test was then done to ascertain the significance of the differences between the percentages of the two patient groups on each characteristic, (2) the method of sorting by 4 experienced judges,[2] and (3) a comparison of the mean number of figures recalled and drawn from memory by each group of patients. The *t* test was also used here.

The results of the comparisons were largely negative. Among all the comparisons made, the only significant differences found were those between the two groups on the characteristics "erasing of figures" and "retracing of figures". The functional group exhibited significantly fewer of both of these tendencies; for the first characteristic the *t* equals 2.10 with a probability of less than .05, while for the second the *t* equals 3.62 with a probability of less than .001.

Perhaps the most interesting finding was that each judge separately, and all four judges combined, were unable to differentiate the protocols of the two types of patients in a manner significantly different from chance. This is the acid test with a technique like the Bender Gestalt where there is no standard "objective" scoring system. In view of the frequent use of this test with neurotic patients, more studies of this type seem indicated, perhaps utilizing different methods of evaluating the protocols and/or involving different catagories of patients.

---

*Based on research done by the author at the Minneapolis Veterans Administration Hospital. The opinions expressed and conclusions drawn are those of the author and they do not necessarily reflect the viewpoint of the Veterans Administration.

[1] Some samples of the characteristics rated are: rotation of figures, embellishment, perseveration, changes in size, etc.

[2] The chi-square test was used in this comparison.

WRITING PSYCHOLOGICAL REPORTS

AUSTIN FOSTER

University of Texas

Most students in clinical psychology are poorly trained in report writing techniques. This is almost entirely due to their instructors, who, being psychologists themselves, accept technical test terminology as proper, and, being teachers, expect complete and detailed reports. In a mental hospital, one writes psychological reports for a *psychiatrist*. The report should be designed to supply the information desired by the referring physician in language which is meaningful in terms of the patient, not in a kind of test cant. One of the principal justifications for a psychological program is time-saving. If testing cannot elicit information in less time than a series of clinical interviews, there is little justification for the program. Students who are trained to give a tremendous number of tests, and to take four or five hours to score and interpret the results are mistrained. To be of value to a psychopathic hospital, the psychologist (if testing is his primary function) should be able to test two patients a day, including all interpretation and report writing.

The psychologist should be able to report the results of his work without wind and water. Students are frequently taught to pad their reports with inane behavioral observations, irrelevant details of the testing, and fine-sounding writing with less in it than meets the eye. Actually, it is the rare report which requires more than one typewritten page. The following guides will make for better reports:

1. Include no test shop talk. Phrases like co-arctation, contamination, digit-span/arithmetic ratio, or position responses indicate that the writer is unable to interpret his data, and is leaving that task to the psychiatrist. Unless the psychiatrist is also a psychologist, this is the greatest sin.

2. Do not list the tests and discuss the results test by test. The psychologist should be the judge as to which tests are to be used; it is assumed that he has sufficient sense to do an adequate testing job. If the psychologist cannot integrate his results into a single interpretation of the test battery, he is inadequately trained.

3. Be positive when sure, qualify when in doubt. Some psychologists appear to believe that it increases respectability to preface every phrase with a qualifiier such as "it appears" or "the data suggest" or "one may speculate." The psychiatrist has a right to know when one is on firm ground and when one is merely suggesting a possibility. The psychiatrist will treat the report as fallible anyway, just as he would any other laboratory report.

4. Occasionally there will be a test record that just doesn't make sense. Admit it. No report at all is better than any examiner's fantasy.

EDITORIAL OPINION

COMMENTS ON THE DEVELOPMENT OF THE PSYCHOANALYTIC MOVEMENT

It is paradoxical that the psychoanalytic movement is attaining new peaks of recognition and acceptance at a time when scientific opinion increasingly questions its validity. Many professional groups have espoused psychoanalytic theory (in some cases in direct opposition to Behavioristic viewpoints formerly held), and it is currently fashionable to follow analytic thinking which is now considered to be the most dynamic viewpoint. However, in spite of the apparent acceptance of psychoanalysis by many leading "authorities" and in the best circles, it is still extremely difficult to obtain any objective evidence concerning the validity of its teachings. So many claims are being made advocating the value of psychoanalysis that it seems important to offer a critical evaluation of some aspects of the movement which should give pause to an overenthusiastic and uncritical acceptance of the viewpoint among clinical psychologists whose experimental-statistical basic science training should make them less prone to accept the unproven. If the psychoanalytic movement is ever to attain true scientific respectability it will be necessary to modify some of the trends described below and which may be observed on the current psychoanalytic scene.

1. In some respects the psychoanalytic movement appears to be bypassing experimental psychological science and modern medicine. Although a few analytic training centers have been established in close proximity to graduate psychology departments and medical schools, this proximity is geographic rather than functional. There appears to be a tendency for the analysts to insulate themselves away from the main stream of scientific progress. There appears to be insufficient effort to integrate the data of objective science with analytic theory and clinical practice. There is evidence that many psychoanalytic training centers and practitioners are oblivious to developments in basic psychological science. At least, there is little evidence that analysts-in-training are being systematically exposed to basic training in psychological science.

2. The psychoanalytic movement has not made systematic use of experimental and statistical research methods which are standard practice in basic psychology. Most psychoanalytic studies still depend heavily on the anecdotal method supported only by metaphysical theorizing and personal opinion. One looks largely in vain for experimental-statistical studies, and it must therefore be concluded that the scientific validity of many psychoanalytic theories has not yet been established. Is the first allegiance of the analyst to the scientific method in general or to his own discipline?

3. To many, the promotional development of the psychoanalytic movement has not been scientifically healthy. Many analysts show intolerant attitudes toward non-analytically trained personnel, i.e. "I don't see how anyone can do therapy unless he has been analyzed." There is frequently encountered an either-or attitude of "either you get psychoanalyzed or you will not be considered competent professionally." The mere fact of having been analyzed has become accepted almost as a badge of competency; one is either a member of the favored few or an outsider. Such attitudes are completely unscientific and reflect the failure to place proper emphasis on basic science and standardization of training as the best qualifications for true professional competency.

4. The multiplication of "schools" of psychoanalysis appears to be an unhealthy development reflecting ideological inconsistencies in the movement. At a time when basic science is attempting to achieve worldwide uniformity and standardization of subject matter, the analytic movement has splintered with the appearance of new schools which are frequently highly competitive. The loyal orthodox group has striven to close up the ranks by insisting upon rigid adherence to the Gospel according to Freud but even this has been a disadvantage since it has inhibited healthy evolution of his theories. The spectacle of the psychoanalytic movement seeking to close schisms by voting on acceptable versions of its theory while rejecting apochryphal ones bears more resemblance to theology than to science.

5. The language of psychoanalysis appears to be semantically invalid in many respects. Psychoanalytic terminology has become almost a "lingo" with the use of many ill-defined concepts and fuzzy abstractions which are too undifferentiated to permit of objective verification. The birth

of schools of neopsychoanalysis has resulted only in increasing the complexity of semantic confusion. Probably psychoanalysis can never be scientifically validated until it objectifies its vocabulary and theoretical concepts.

6. Psychoanalytic preoccupation with the symbolism of the Unconscious may be an unfortunate critical development. Preoccupation with symbolism has led to the attempt to read unconscious significance into all behavior. Without having made any objective studies to determine whether any particular act *has* or *has not* any symbolic significance, analysts tend to point a *markedly accusing* finger at almost all behavior by giving it morbid (usually psychoneurotic) significance. No act, however innocent, escapes the attempt to see sinister meanings behind it. This tendency involves the error of projecting the analyst's own interpretations on behavior.

7. Psychoanalytic doctrines have been uncritically accepted without sufficient effort to determine their validity. The present scene is characterized by the widespread use of psychoanalytic concepts by all types of professional personnel and by the general public without any critical evaluation of what is involved. Scientific psychology held out against this trend for a long time, but many clinical psychologists have swallowed these concepts whole probably because of their professional proximity to analytically oriented psychiatrists upon whom they have had to depend for protection and support.

8. The methods of some psychoanalytic training centers appear to verge on mysticism and cultism. In some localities, the matter of the training analysis has almost assumed the proportions of a racket. In order to enter the ranks of the psychoanalytic "elite", one must secure the training analysis from the "right" persons, preferably a disciple of Freud the Master. But then a training analysis is not enough; one must also obtain "therapeutic analyses" for wife and children. Persons simultaneously taking analysis from one analyst become "siblings in analysis" and presumably experience all the conflicts of the family hierarchy. A large source of the revenues of some analysts appears to come from pupils seeking to enter the field.

9. The prestige value of being a psychoanalyst has been overexploited in some centers to the detriment of basic psychological science as a whole. In some places, honors and promotions are available only to recognized psychoanalysts. In order to be accepted into the fold, young students must cater to authority and dogma. Personnel without analytic training sometimes find themselves discriminated against.

10. Psychoanalysis has occasionally been advertised to the general public in ways which border on the unethical. In the last twenty years there have appeared a large number of books written by prominent psychoanalysts in which unsupported claims have been made for psychoanalysis which go far beyond the range of what has been scientifically validated. Such books have catchy titles, sensational style, and contain frank advertising both for the movement and for individual analysts. Much of the material included in these books consists of dialectic exercises in which the imagination is allowed to soar unfettered by the limitations imposed by scientific methods.

11. There is little evidence that psychoanalysts as a group are any more "normal" and free of neurotic behavior in spite of their having been exposed to intensive analytic treatment and training. Detailed observation of the personal lives of psychoanalysts fails to reveal that they have achieved any Olympian detachment from human weaknesses in spite of their proprietorship of the Truth. They appear to be neither more nor less subject to behavior problems than any other comparable professional group. If anything, student selection methods in psychoanalysis have been less rigid than in comparable disciplines.

12. In the last analysis, the most scathing criticism of psychoanalysis is that today, after more than 50 years of development, most psychoanalytic doctrines are still unproven by the most modern experimental or statistical methods. At its best, psychoanalysis has contributed some penetrating insights. At its worst, some of the excesses perpetrated in the name of psychoanalysis have been so extreme that they dwarf the shortcomings of the older systems which psychoanalysis seeks to supersede. Viewed in terms of the relentlessly critical attitudes characteristic of all true scientific method, it must be concluded that psychoanalysis is not science but only extremely brilliant, and sometimes extremely erroneous, conjecture.

All this is not intended to involve any blanket condemnation of modern psychoanalysis, but simply to underline some developments which seem unhealthy. No one can deny the contribution made by Freud and psychoanalysis, particularly in uncovering a wealth of clinical case materials formerly ignored and in seeking to discover the mental dynamisms underlying them. But failure to view the total problem objectively has the obvious dangers of subtly blocking the evolution of objective basic science. As in war where the basic issue boils down to the question of whether you are with us or against us, so in science we must constantly reexamine our consciences in order to determine if we are being totally and consistently scientific in our attitudes.

F. C. T.

BOOKS RECEIVED

MOWRER, O. H. *Learning Theory and Personality Dynamics.* New York: Ronald, 1950. $7.50.

DOLLARD, J. AND MILLER, N. E. *Personality and Psychotherapy.* New York: McGraw-Hill, 1950. $6.50.

CAMERON, D. E. *General Psychotherapy.* New York: Grune and Stratton, 1950. $5.00.

GULLIKSEN, H. *Theory of Mental Tests.* New York: Wiley, 1950. $6.00.

ABT, L. E. and BELLAK, L. (Eds.) *Projective Psychology.* New York: Knopf, 1950. $6.00.

HAHN, M. E. *General Clinical Counseling.* New York: McGraw-Hill, 1950. $3.50.

ROBINSON, F. P. *Principles and Procedures in Student Counseling.* New York: Harper, 1950. $3.00.

CATTELL, R. B. *Personality.* New York: McGraw-Hill, 1950. $5.50.

BRAYFIELD, A. H. (Ed.) *Readings in Modern Methods of Counseling.* New York: Appleton-Century-Crofts, 1950.

MIKESELL, W. H. *Modern Abnormal Psychology.* New York: Philosophical Library, 1950. $10.00.

COUTU, W. *Emergent Human Nature.* New York: Knopf, 1949. $3.75.

FROMM-REICHMANN, F. *Principles of Intensive Psychotherapy.* Chicago: Univ. Chicago Press, 1950. $3.75.

RAPAPORT, D. *Emotions and Memory.* New York: International Universities Press. 1950. $4.00.

BLEULER, E. *Dementia Praecox.* New York: International Universities Press, 1950. $7.50.

WORTIS, J. *Soviet Psychiatry.* Baltimore: Williams and Wilkins, 1950. $5.00.

POLLAK, O. *The Criminality of Women.* Philadelphia: University of Pennsylvania Press, 1950. $3.50.

KLEIN, H. R., POTTER, H. W., and DYK, R. B. *Anxiety in Pregnancy and Childbirth.* New York: Hoeber, 1950. $2.75.

HARROWER, M. R. (Ed.) *Recent Advances in Diagnostic Psychological Testing.* Springfield, Ill.: Thomas, 1950. $2.75.

WOLFF, W. *Values and Personality.* New York: Grune & Stratton, 1950. $4.75.

ANDERSON, C. M. *Saints, Sinners and Psychiatry.* Philadelphia: Lippincott, 1950. $2.95.

GRECO, M. *Group Life.* New York: Philosophical Library, 1950. $4.75.

HADFIELD, J. A. *Psychology and Mental Health.* New York: Macmillan, 1950. $2.75.

NAUMBERG, M. *Schizophrenic Art: Its Meaning in Psychotherapy.* New York: Grune & Stratton, 1950. $10.00.

Various Editors. *The Psychoanalytic Study of the Child, Vol. V.* New York: International Universities Press, 1950. $7.50.

MORGAN, C. T. and STELLAR, E. *Physiological Psychology, 2nd Ed.* New York: McGraw-Hill, 1950. $5.00.

BETTELHEIM, B. *Love is not enough.* Glencoe, Ill.: The Free Press, 1950. $4.50.

LANDIS, C. and BOLLES, M. M. *Textbook of Abnormal Psychology.* New York: Macmillan, 1950. $5.00.

FISHER, V. E. *The Meaning and Practice of Psychotherapy.* New York: Macmillan, 1950. $5.00.

DIETHELM, O. *Treatment in Psychiatry, 2nd Ed.* Springfield, Ill.: Thomas, 1950. $8.50.

GREENBLATT, M., ARNOT, R. and SOLOMON, H. C. (Eds.) *Studies in Lobotomy.* New York: Grune & Stratton, 1950. $10.00.

DONAHUE, W. and DABELSTEIN, D. *Psychological Diagnosis and Counseling of the Adult Blind.* New York: American Foundation for the Blind, 1950. $2.00.

DORCUS, R. M. and SHAFFER, G. W. *Textbook of Abnormal Psychology.* Baltimore: Williams and Wilkins, 1950. $5.00.

FISCHER, S. *Principles of General Psychopathology.* New York: Philosophical Library, 1950. $4.75.

MYKLEBUST, H. *Your Deaf Child.* Springfield, Ill.: Thomas, 1950. $2.50.

ENGEL, G. L. *Fainting.* Springfield, Ill.: Thomas, 1950. $2.75.

PORTEUS, S. D. *The Porteus Maze Test and Intelligence.* Palo Alto, Calif.: Pacific Books, 1950. $4.00.

ROHEIM, G. *Psychoanalysis and Anthropology.* New York: International Universities Press, 1950. $10.00.

FODOR, N. and GAYNOR, F. *Freud: Dictionary of Psychoanalysis.* New York: Philosophical Library, 1950. $3.75.

Various authors. *Current Trends in the Relation of Psychology to Medicine.* Pittsburgh: University of Pittsburgh Press, 1950. $3.75.

SCHILDER, P. *The Image and Appearance of the Human Body.* New York: International Universities Press, 1950. $4.50.

CRAWFORD, P. L., MALAMUD, D. I. and DUMPSON, J. R. *Working with Teen-Age Gangs.* New York: Welfare Council of New York City, 1950.

HERON, W. T. *Clinical Applications of Suggestion and Hypnosis.* Springfield, Ill.: Thomas, 1950. $3.00.

BACK IN STOCK!

ENCYCLOPEDIA OF PSYCHOLOGY

Edited by Philip Lawrence Harriman, Ph.D.
with 80 Eminent Collaborators

This monumental work presents an up-to-date and complete survey of all the major issues and concepts in the field of psychology. The 900 pages contain almost 100 concepts and movements treated by 80 eminent contributors—psychologists, psychiatrists, psychoanalysts and educators. These articles deal with all major topics in modern psychology.

Great care was taken in giving correct and objective interpretations of the various theories, schools and terms germane to the subject. Whenever possible, the issues at hand were dealt with in a manner that would explain controversial problems, not one-sidedly, but from all accepted points of view.

Much of the psychological experimental work of our time is described here by its representatives or by those who had occasion to observe these studies intimately.

These sections are arranged alphabetically, each supplemented by bibliographical references.

"An ably edited and well presented reference work . . . it covers not only general psychology but dynamic psychology, psychopathology, and the mental disorders. The 80 collaborators are well known in their fields."

—The Psychiatric Quarterly.

900 pages $5.00

Special student bulk rate on 10 or more. Expedite shipment by prepayment.

PHILOSOPHICAL LIBRARY, *Publishers*
15 East 40th Street, Desk 190 New York 16, N. Y.

PRINCIPLES OF PERSONALITY COUNSELING

by FREDERICK C. THORNE, M.D.
University of Vermont

This new textbook presents a new and eclectic system of Personality Counseling, including detailed discussions and case presentations of all known methods.

PRICE: $5.00. ORDER NOW

THE MONOGRAPH SERIES

No. 1.	Billingslea: Bender Gestalt Test	$ .90
No. 2.	Carter & Bowles: Qualitative Aspects of Psychological Examining	$1.25
No. 3.	Training in Clinical Psychology	$1.50
No. 4.	Critique of Nondirective Therapy	$1.25
No. 5.	Buck: The H-T-P Manual	$2.00
No. 6.	Karpman: Objective Psychotherapy	$2.00
No. 7.	Statistics for the Clinician	$2.00
No. 8.	Tyson: Current Mental Hygiene Practice	$2.00

ORDER TODAY

Journal of Clinical Psychology

BRANDON, VERMONT

CLINICAL STAFF

MEDICAL STAFF OF PENNSYLVANIA

Leslie R. Angus, M.D.
Robert Devereux, M.D.
Ruth E. Duffy, M.D.
Herbert H. Herskovitz, M.D.
Ruth Stephenson, M.D.

PSYCHOLOGICAL STAFF OF PENNSYLVANIA

Edgar A. Doll, Ph.D.
Director of Research
Michael B. Dunn, A.M.
Robert G. Ferguson, A.M.
Edward L. French, Ph.D.
Mary J. Pawling, A.M.

MEDICAL STAFF OF CALIFORNIA

Charles M. Campbell, Jr., M.D.
Director of Clincal Services
Richard H. Lambert, M.D.
Consulting Psychiatrist
Ivan A. McGuire, M.D.
Consulting Psychiatrist
David L. Reeves, M.D.
Consulting Neurologist

Thomas W. Jefferson, Ph.D.
Clinical Psychologist

STABILITY
through Education with Therapy

Emotional growth must keep pace with intellectual development if the child is to become a stable, well-adjusted adult . . . successful, happy and valuable to society. The disturbed child is doubly handicapped — emotional imbalance interferes with his acceptance of education and carries its own burden of character maladjustment.

When you meet the problem of providing a school-age patient with specialized guidance, we invite you to let us evaluate the needs. When the intelligence is normal but emotional disturbances block the ability to learn, our experienced staff will carefully study the details and offer a considered report.

Please address your inquiries to
John M. Barclay, Registrar

DEVEREUX FOUNDATION
CALIFORNIA ..... PENNSYLVANIA

Devereux Schools

HELENA T. DEVEREUX, Director
SANTA BARBARA, CALIFORNIA :-: DEVON, PENNSYLVANIA